Essays in Market Design and Industrial Organization

DISSERTATION

zur Erlangung des akademischen Grades
doctor rerum politicarum
(Doktor der Wirtschaftswissenschaft)

eingereicht an der
Wirtschaftswissenschaftlichen Fakultät
der Humboldt-Universität zu Berlin

von
Philipp D. Dimakopoulos, M.Sc.

Präsidentin der Humboldt-Universität zu Berlin:
Prof. Dr.-Ing. Dr. Sabine Kunst

Dekan der Wirtschaftswissenschaftlichen Fakultät:
Prof. Dr. Christian D. Schade

Gutachter:
1. Prof. Dr. Roland Strausz
2. Prof. Dr. Ulrich Kamecke

Eingereicht am: 21.12.2017
Tag des Kolloquiums: 13.04.2018

To my parents.

Acknowledgements

I am very grateful to my supervisor Roland Strausz, whose guidance and support throughout my time as a PhD student have been invaluable. I would also like to sincerely thank Ulrich Kamecke for his advice and support.

Moreover, I am grateful to Pio Baake and Helmut Bester for all their insightful comments and their guidance. Furthermore, I want to thank Dorothea Kübler and Morimitsu Kurino for their suggestions and for introducing me to the topic of matching market design. I would also like to thank Ulrich Doraszelski for his comments and for introducing me to the field of computational industrial organization. Further, I thank Dirk Engelmann for his guidance at the Chair.

Without my coauthors this thesis would not be possible. C.-Philipp Heller worked with me on chapter one and Slobodan Sudaric worked with me on chapter two.

I am thankful to many fellow PhD students and postdocs in Berlin for all the valuable feedback, and especially to my office mates Andreas Asseyer, Thomas Schacherer and Jonas von Wangenheim, also for the lively and fun discussions.

For their administrative support, I thank Regine Hallmann, Viviana Lalli, Myrna Selling and Sandra Uzman.

I wrote this thesis while I was at Humboldt-Universität zu Berlin, but I also spent time at the Freie Universität Berlin and at the Ludwig-Maximilians-Universität München, for which I am thankful to Florian Englmaier.

I would like to thank the participants of the Microeconomic Theory Colloquium in Berlin for all the discussions at various stages of the three chapters. For their comments on chapter one I would like to thank Christian Basteck, Lucien Frys, Maciej Kotowski, and participants of the 12th Meeting of the Society for Social Choice and Welfare in Boston, the Jahrestagung of the Verein für Socialpolitik in Münster, as well as two anonymous referees. For their comments on chapter two I would like to thank Colin von Negenborn and Markus Reisinger, as well as participants of the IODE in Liège, BCCP in Berlin and EARIE in Maastricht, and two anonymous referees. For their comments on chapter three I would like to thank Alesandro Bonatti, Thomas Faggart and Vincent Meisner, as well as participants at the CRESSE in Rhodes and the EARIE in Lisbon.

Financial support from the DFG is gratefully acknowledged.

Contents

List of Figures

Introduction

In the three independent chapters of this thesis I study microeconomic interactions on different markets, thereby addressing questions of market design and industrial organization. Although distinct in their nature, all markets I consider allow for the interplay and competition of participants, whose decisions impact the allocation of limited resources.

In this work market design refers to the design of allocation rules affecting the outcome in two-sided matching markets, as considered in chapter one. Here, participants on either side have preferences over a matching with the other side. While there is competition among participants for these matchings, there is no price to balance supply and demand. Therefore matching algorithms are employed to find an allocation. When there are prices to clear a market in the traditional microeconomic sense, as in chapters two and three, tools from game theory and industrial organization help to analyze the strategic behavior of market participants and to understand what makes these markets work better. However, in online markets as in chapter two, prices may not be monetary but rather consist of private data "payments", which has important implications for privacy and competition policies. Moreover, on many other markets as in chapter three, prices can be dynamic because goods are sold during longer sales periods, such that inter-temporal optimizations of firms and consumers need to be considered. Summing up, as the notions of market and price can differ, also analyses and policy results in the three considered markets are distinct.

In the first chapter of this thesis I consider the market for lawyer trainee-ship positions, which exhibits a many-to-one matching problem. Based on lawyers' preferences and their priorities at courts, many lawyers are matched each year to a regional court in Germany. Because of excess demand in some regions lawyers often have to wait before being allocated. I show that the currently used "Berlin" mechanism is not weakly Pareto efficient, does not eliminate justified envy and does not respect improvements. Therefore, I introduce a mechanism based on the matching with contracts literature, using waiting time as the contractual term. The resulting mechanism is strategy-proof, weakly Pareto efficient, eliminates justified envy and respects improvements. Furthermore, I extend the proposed mechanism to allow for a more flexible allocation of positions over time.

In chapter two I analyze online platform competition. Here, two-sided platforms, such as search engines or social networks, charge on the one side a monetary price from advertisers for the placement of ads, while on the other side users "pay" with their private data in order to gain access to the platform. This user data improves ad-targeting, hence is a valuable resource for platforms and advertisers. Considering that users incur privacy costs, I show that the equilibrium level of data provision is distorted and can be inefficiently high or low: if overall competition is weak or if targeting benefits are low, too much private data is collected, and vice-versa. Further, I find that softer competition on either market side leads to more data collection, which implies substitutability between competition policy measures on both market sides. Moreover, if platforms engage in two-sided pricing, i.e. use monetary payments on both market sides, data provision would be efficient.

The third chapter of this thesis is on dynamic pricing, such as in markets for airline tickets or travel bookings. Competition in these markets does not only take place in a static environment but rather throughout a finite selling horizon with a deadline, for example the departure or event day. Capacity-constrained firms compete to sell their goods, whereby they can dynamically adjust their prices at all times until this deadline. Therefore, firms face an inter-temporal pricing problem, trading off current-period payoffs and continuation values. Forward-looking consumers, too, face an inter-temporal problem of buying at current prices or waiting to buy later at possibly better prices, however at the risk of being rationed. In the pure-strategy equilibrium firms set dispersed prices, which depend on the number of unsold capacities relative to remaining selling time. The resulting price paths provide an explanation for empirically observed price volatility. A policy allowing consumers to become forward-looking increases consumer surplus yet reduces efficiency and industry profits. Further, I find that competition policy is especially relevant if market capacities are excessive. Last, my results show that ex-ante capacity production can be inefficiently low.

Chapter I

Matching with Waiting Times: The German Entry-Level Labor Market for Lawyers

Based on Dimakopoulos and Heller (2017).

1 Introduction

Many real world matching markets fail to match all participants. Those who are unmatched may either leave the market altogether or wait and participate in a later matching procedure. The example that we study here is the allocation of graduating lawyers to their legal trainee-ship at regional courts in Germany. In this market congestion arises because of excess demand for positions in some parts of the country. This congestion is managed by requiring unmatched applicants to enter a wait list for their trainee-ship. To ensure that lawyers will eventually obtain a position at a court, the priority of a lawyer increases with the acquired waiting time. We assume that lawyers have preferences over where and when they complete their legal trainee-ship. The preferences over time are however ignored by the currently used procedure, which leads to a lack of efficiency, justified envy and a lack of respect of improvements. We propose a new procedure that does not suffer from those shortcomings.

The focus of this work is the trainee-ship allocation problem between graduated lawyers on the one side and courts on the other side. This is an important market as each year there are approximately 8,000 positions for legal trainee-ship in Germany.[1] These numbers are comparable to the (roughly) 20,000 US hospital residency program

[1]Based on authors' calculations using data from `http://www.juristenkoffer.de/rechtsreferendariat/` (Accessed 8. October 2015).

matches per year studied by Roth and others, e.g. in Roth (1984). In this lawyer market the wage is regulated so it cannot be used to reduce congestion by balancing excess demand.

Unlike in the United States, in Germany lawyers typically begin their legal education as an undergraduate, studying law at a university for around four years. Afterward students take a first state exam, set by the 16 federal states. Following this, students may apply for a legal trainee-ship. Completion of the trainee-ship is necessary to practice law in Germany and a requirement for many jobs in the bureaucracy. It is thus important to ensure access to trainee-ships for all lawyers who wish to complete it.

There is no cooperation across the federal states in terms of having a single national market for positions. This means that each lawyer can in principle apply for a position in each of the federal states.[2] This leads to sizable congestion, since in the extreme the total number of registered applicants in the system will be several times the number of positions actually demanded by the lawyers. We suspect that multiple applications by some lawyers for positions in several federal states is at least partially responsible for long waiting times as some lawyers apply for positions as a safety option which they are unlikely to take. The authorities seem to be aware of this possibility. Several application forms contain declarations that there are no pending applications to other regional courts of appeals or require applicants to withdraw other applications.[3] If they do not accept a position after they have been offered one, then the application system requires the lawyers to inform the respective authorities in the federal state.[4] The authorities may then decide to allow additional lawyers to begin their trainee-ship at that period. This process of refusing and making new offers takes up time and may leave some positions unused if no other lawyers can be found to take up these positions.[5] Note that this process of formally accepting and rejecting offers could also be addressed by allowing

[2]The allocation of lawyers to courts is organized by regional courts of appeals in each state. Some federal states, notably Bavaria and North Rhine-Westfalia, contain several regional courts of appeals so even within a state there is scope for greater coordination.

[3]See for example the application web page of the regional court of appeals in Munich, `https://www.justiz.bayern.de/gericht/olg/m/studiosi/01441/index.php` (accessed 8. October 2015).

[4]For example, Art. 4 of the "Verordnung über die Aufnahme in den juristischen Vorbereitungsdienst" of Hamburg states that applicants who have not accepted a position that was offered to them within 10 days, will not be allocated. Furthermore it says that if an applicant does not accept a position twice, the applicant will be excluded from the application procedure and will have to reapply. Last, it says positions which have not been accepted are allocated to applicants next in line.

[5]For example on 6. October 2015 in North Rhine-Westfalia there were 12 positions still to be filled to begin on 1. November 2015. See `http://www.olg-duesseldorf.nrw.de/aufgaben/referendarabteilung/09_weiter_info/index.php` (accessed 8. October 2015). Congestion problems arising from the need to sequentially inquire about agents acceptance and rejection of offers have been found for example in the market for clinical psychology, Roth and Xing (1997).

lawyers to formally declare some courts as unacceptable and committing applicants to accepting any position that they were offered. However if the value of the outside option is unknown at the time preferences are submitted, then this will not alleviate the problem of offers being refused.[6]

While the organization of the allocation procedure by federal state, rather than having a national procedure, seems to us to be a major cause of congestion in the market for lawyers, addressing this problem would require coordinated action of the federal states. Given the difficulty of establishing such cooperation, we here take the approach to consider only an isolated federal state and analyze how the allocation procedure in that federal state should be designed to better handle the congestion resulting from the lack of national coordination. We thus treat the federalized nature of this labor market as an additional constraint to be respected by the market designer, akin to the constraint that monetary flexibility cannot be used to clear some matching markets (e.g. for kidneys or schools[7]).

The number of available positions for the trainee-ship varies by court and usually depends on its size and the budget that has been made available for legal trainees in the budget of the federal state and/or the capacity of the court. This budget is usually set for several starting dates in advance. For example, the relevant administrative order in Berlin states that the capacity needs to be determined for one year in advance while positions may be started in February, May, August and December.[8] In Hamburg, the relevant administrative order states that the number of positions is determined by the number of positions fixed in the budget, which is valid for at least one year. Trainee-ships can start every even-numbered month.[9] In Hessen capacities are set every half a year, while new positions are available in all odd-numbered months.[10]

Due to large numbers of applications in some federal states, not all lawyers applying for a position at a court can be allocated at their desired starting time.[11] The excess

[6] This is the case when several federal states run their allocation procedures in parallel.

[7] See Roth et al. (2004) and Abdulkadiroğlu and Sönmez (2003)

[8] See Art. 2 (2) in the "Verordnung über die Ausbildungskapazität und das Vergabeverfahren für den juristischen Vorbereitungsdienst" (Stadt Berlin (2004)) for setting capacities. Dates for entry into the trainee-ship are taken from `http://www.berlin.de/sen/justiz/gerichte/kg/ausbildung/jur-vorb/bew-verf/` (accessed 6. October 2015).

[9] See Art. 2 and Art. 3 (1) in the "Verordnung über die Aufnahme in den juristischen Vorbereitungsdienst" (Hansestadt Hamburg (2012)) for how capacities are set and the dates when trainee-ships start, respectively.

[10] See the guidelines on the legal trainee-ship for Hesse.

[11] Most application forms ask for the desired entry date of an applicant. Even if applicants could only apply for the next starting date, delaying applications until that date ensures that students can affect the time period for which they are considered.

demand is managed via a system based on waiting times accumulated by the applicants.[12] Most federal states have a system whereby a lawyer's priority in being allocated a place at a court increases in the number of times that lawyer was not matched. For example in Hamburg, grades, waiting time and other concerns are weighted and expressed as a single score for each lawyer.[13] In North Rhine-Westfalia by contrast only the time since the application was received by the regional court of appeals determines the ranking of a candidate.[14] In Hessen 35% of positions are reserved for applicants with the highest waiting time.[15] In Brandenburg 70% of positions are reserved for applicants with the highest waiting time.[16] Thereby it is in principle possible for each lawyer to gain some place in a federal state eventually. Currently, average waiting times can be up to 24 months, depending on the federal state, although it should be noted that in many states waiting time is zero or only a couple of months.[17]

When applying for a position lawyers can typically indicate a preference for a particular regional court.[18] While lawyers can submit rankings over the courts, there is no legal guarantee of being assigned the first choice court.[19] While in general the allocation of lawyers to courts should take into account reported preferences, capacities and priorities, we could not find a clear description of the methods used to allocate lawyers to

[12]For example, in Berlin for entry on August 3rd 2015 applicants with a grade of 10 or higher were admitted if they applied 5 months earlier. Those who did their state exam in Berlin were admitted if they applied 10 months earlier, while those who did their state exam elsewhere with a grade below 10 were admitted if they applied 11 months earlier.

[13]Art. 5 of the Aufnahmeverordnung (AVO,Hansestadt Hamburg (2012)) sets rules on how to calculate this score. The base score is the minimum of 6.49 and the grade achieved by the lawyer in the first state exam (Art. 5 (1) AVO). Further points can be added for example for having completed military service, disabilities, having done the state exam in Hamburg and for every 6 months of accumulated waiting time (Art. 5 (2) AVO). In case of ties in the weighted score, Art. 6 (1) AVO instructs to use the grade in the state exam to break ties. Remaining ties are to be broken via lottery according to Art. 6 (2) AVO.

[14]See `https://www.justiz.nrw.de/WebPortal/JM/landesjustizpruefungsamt/juristischer_vorbereitungsdienst/2Einstellung/index.php` (accessed 7. October 2015).

[15]See the Justizprüfungsamt Hessen (2011). Another 50% are reserved for lawyers based on merit and the remaining 15% are reserved for applicants satisfying social criteria.

[16]See Art. 11 (3) of the "Juristenausbildungsgesetz" of Brandenburg (Land Brandenburg (2014)). 20% of positions are given based on waiting time, with the remaining 10% given based on social criteria.

[17]Based on data from `http://www.juristenkoffer.de/rechtsreferendariat/` (Accessed 8. October 2015).

[18]For example lawyers applying to do their trainee-ship in the district of the Dusseldorf (North Rhine-Westfalia) regional court of appeals can apply to the regional courts in Dusseldorf, Duisburg, Kleve, Krefeld, Mönchengladbach or Wuppertal.

[19]For example, Art. 30 (3) of the Lawyer Education Law of North Rhine-Westfalia (Juristenausbildungsgesetz Nordrhein-Westfalen, JAG NRW) states that there is no legal right to a position in a particular district of a regional court of appeals and at a particular time.

courts.[20] Some regional courts of appeal give some additional insights into how lawyers are allocated to particular courts. For example, applicants to Munich are ranked according to a number of criteria.[21] The highest priority is given to applicants having to care for their children, followed by married couples having lived for at least one year in the desired location. Next come those suffering from serious illnesses and then those working as teaching assistants at universities in the desired location. Finally, the length of time that applicants have lived in the desired location is used. There is however no indication in what way those priorities are used.

To analyze the market while accounting for waiting time, we propose a lawyer-court matching problem based on Hatfield and Milgrom (2005).[22] On the one side of the market there are lawyers, who have preferences over assignments to courts over time. Courts on the other side have priorities over lawyers, possibly based on their grade, social criteria and accumulated waiting time, which together with the current time period determines a lawyer's waiting time.[23] A matching mechanism in this context produces an allocation consisting of a subset of contracts, which specify a lawyer, a court and the time period the trainee-ship begins. Capacities of a court in future periods are already known, as we discussed above.

Based on the features of the currently used procedure we introduce the "Berlin" mechanism. This mechanism is not weakly Pareto efficient. We show that this mechanism may lead to allocations where one lawyer justifiably envies another. Furthermore improvements of the ranking achieved by a lawyer may yield an allocation that is worse for that lawyer. However, by construction, the Berlin mechanism achieves an allocation, such that no currently available positions remain unfilled while allocating some lawyers to later positions.

We propose the time-specific choice function, which is a special case of choice functions based on slot-specific priorities of Kominers and Sönmez (2016). Here time-specific means that each court can only accept a fixed number of students to begin their trainee-ship in a given period. Using the time-specific choice functions, the cumulative offer process of Hatfield and Milgrom (2005) is used to find stable allocations. Extending beyond current

[20]For example, in the guidelines on the application in the Dusseldorf district, it simply says that lawyers are allocated to courts following a "comprehensive view" of all applications. This may result in lawyers not getting their first choice so that they are asked to indicate further preferences, (Oberlandesgericht Düsseldorf (2015b)).

[21]See the criteria for the allocation of trainee-ships, Oberlandesgericht München, (2015).

[22] Other related papers are Hatfield and Kojima (2010), Kominers and Sönmez (2016), Sönmez (2013) and Sönmez and Switzer (2013).

[23]In the district of the regional court of appeals in Dusseldorf, it is explicitly stated that a higher waiting time does not affect the allocation to a desired court (Oberlandesgericht Düsseldorf (2015b)).

results, we can show the existence of a lawyer-optimal stable allocation, when lawyers prefer earlier assignments. In cases where lawyers' preferences are unrestricted, no such lawyer-optimal stable allocation need exist.

The time-specific choice function does not satisfy some properties used in the previous literature. Notably it fails to satisfy the unilateral substitutes condition and the law of aggregate demand. Hence we cannot use the results of Hatfield and Kojima (2009) and Hatfield and Kojima (2010). We instead apply the results of Kominers and Sönmez (2016) to show that the time-specific lawyer proposing mechanism is (group) strategy-proof for the lawyers. Moreover, this mechanism is weakly Pareto efficient, eliminates justified envy and respects improvements. Furthermore, our mechanism creates incentives for all lawyers to report verifiable information increasing their priority at a court. However, it may allocate some lawyers to later positions while leaving some currently available positions unfilled. It thus allows current lawyers to obtain better positions at the expense of future lawyers.

We consider another modified version of the matching with contracts model, in which we no longer have time-specific constraints for each court. Instead, courts face only aggregate capacity constraints and are able to shift their positions flexibly over time. This would be applicable if courts had control over their own budgets over a period of some years. We construct the flexible choice function for courts, based on the time-specific choice function. The resulting flexible lawyer-optimal stable mechanism (FLOSM) is (group) strategy-proof, weakly Pareto efficient, while eliminating justified envy and respecting improvements. Furthermore it Pareto dominates the allocation obtained when time-specific capacity constraints need to be respected. It may however violate the time-specific capacity constraints of the courts.

While our model has been developed with the entry-level labor market for lawyers in Germany in mind, there are potentially many more applications of the basic framework. For example, university admissions in Germany for some very competitive courses, such as medicine, often ration places by putting unsuccessful applicants on waiting lists. A certain fraction of all seats is then reserved for those applicants who have waited a sufficient number of periods. Another potential application concerns the allocation of aspiring teachers to teaching trainee-ship positions at schools, in a system very similar to that of lawyers. The main difference to the market for lawyers is that teachers differ based on their chosen subjects, so that schools' preferences over teachers will be more complex than courts' priorities over lawyers. In addition schools are likely to be strategic players, unlike the courts. A position for math and physics teacher could for example be filled either by one teacher for both subjects or by two teachers each responsible for one of the subjects. Further interesting applications of matching with waiting times are (social or student) house allocation problems. For example, if there are a number of different projects to construct social housing that finished at different, known points in time then our model could be directly applied.

The remainder of this paper is organized as follows. In Section 2 we discuss the relevant literature. The model and some definitions are introduced in Section 3. Using our model, in Section 4 we analyze the currently applied Berlin mechanism and its properties. In Section 5 we propose mechanisms based on matching with contracts. Section 7 concludes. The Appendix contains proofs that are not in the main text.

2 Literature

This paper fits into the research agenda started by Gale and Shapley (1962) on two-sided matching. For a summary of research in this vein until 1990, see Roth and Sotomayor (1990). Two-sided matching has found important applications in the design of labor markets. For examples of the application of two-sided matching to medical entry-level labor markets see Roth (1984), Roth (1991) or Roth and Peranson (1999). More recently a number of papers have applied the original two-sided matching problem to the allocation of seats at universities, for instance Balinski and Sönmez (1999) and, more prominently, to the design of school choice mechanisms (Abdulkadiroğlu and Sönmez, 2003; Abdulkadiroğlu et al., 2005b,a).

The canonical model of matching with contracts is due to Hatfield and Milgrom (2005),[24] which was later extended by Hatfield and Kojima (2008), Hatfield and Kojima (2009), Hatfield and Kojima (2010) and Hatfield et al. (2015a). Some early precursors of this type of model are Crawford and Knoer (1981) and Kelso and Crawford (1982).

In two recent contributions by Sönmez and Switzer (2013) and Sönmez (2013) new applications of the matching with contracts model to the allocation of cadets to branches of the US Army are introduced. Their treatment relies heavily on the recent result of the literature on matching with contracts and shows their practical relevance. In their case the number of years a cadet commits to serve in the army is the contract term. In our model the time at which a lawyer starts her trainee-ship is the contract term. Our work is closely related to and makes use of results in Kominers and Sönmez (2016) who study a more general slot-specific matching with contracts model. We use their

[24] Fleiner (2003) uses a similar fixed-point approach to find stable matchings.

results to show (group) strategy-proofness and respect of improvements for our preferred mechanism, the time-specific lawyer offering stable mechanism. We additionally show the existence of a lawyer-optimal stable mechanism by assuming weak impatience for the lawyers.[25] In recent work, Aygun and Turhan (2016) study dynamic reserves in Indian engineering school admission, where some seats might remain unfilled due to affirmative action reserves. Also using the matching with contracts framework and a new choice function for schools, the authors employ privilege types as contractual terms. In our paper we consider the time dimension, giving rise to additional dynamic properties.

There are several papers considering dynamic matching models. Leshno (2015) considers a queuing model in which agents are of two (privately known) types and can be assigned to one of two objects. There is overload in the sense that there are many agents waiting to be assigned an object. This model differs from ours in that the arrival of objects is random, whereas in our model it is known. Furthermore Leshno (2015) assumes that waiting is equally costly for agents, whereas in our model agents differ in time preferences.

Thakral (2015) studies a model similar to Leshno (2015) and ours, where agents are assigned to public housing. In his model houses arrive stochastically over time due to existing tenants moving out of public housing at their discretion. He assumes that agents are weakly impatient in the sense that being assigned public housing earlier is preferred to it being assigned later. Assuming that there is a common ordering of the houses over the agents, he introduces a strategy-proof mechanism that eliminates justified envy and is efficient evaluated at a particular point in time. If the realization of the house arrival process in the model of Thakral (2015) were known, it would correspond to our model in which the available positions in the future are known.

Kadam and Kotowski (2015) consider a two-sided matching model in which agents may have different partners over time. Their model set-up could be formulated in terms of a matching with waiting time model as we propose. The difference to our model is that both sides of the market would be allowed to sign multiple contracts even if all agents can only be matched to one other agent in a given period of time. In addition, they focus on different notions of stability.

Another related literature is the one on dynamic matching markets. Papers in that literature have, to our knowledge, not yet made use of the matching with contracts framework. Damiano and Lam (2005) consider one-to-one matching markets which are repeated over time. Here the outcome is a matching associating one man to a woman for

[25] An alternative route towards our results can be found in Hatfield et al. (2015b) who provide conditions for cumulative offer processes to yield stable and strategy-proof mechanisms.

each period. Similarly, Kurino (2009) considers one-to-one repeated matching markets. The focus in the latter paper is on a new notion of credible group-stable dynamic matchings. The paper by Bloch and Houy (2012) considers the allocation of a set of durable objects to agents who successively arrive and live for two periods. Related, Kurino (2014) considers a dynamic house allocation problem in which agents arrive successively and live for two periods. Abdulkadiroğlu and Loertscher (2007) also consider a dynamic house allocation problem. That paper compares static and dynamic mechanisms, finding that the latter can improve welfare upon the former. Another market design application of dynamic matching problems is Kennes et al. (2014) who consider the allocation of small children to daycare facilities in Denmark. Our paper differs from these papers insofar as in our paper the outcome is a set of contracts in which each lawyer appears only once, so no lawyer is matched repeatedly. Also, unlike the previous papers we make explicit use of the matching with contracts literature, which might also be fruitfully applied in the papers just mentioned. To apply the matching with contracts framework one would simply need to allow lawyers to hold multiple contracts.

This paper is also related to some papers within the theory of matching which analyze different legal entry-level labor markets. Avery et al. (2001) provide empirical data and discuss possible reconstructions of the market for legal clerkships at US federal courts for graduating law students, primarily addressing the unraveling problem. In Avery et al. (2007), the authors describe the unraveling in this market and relate to the problem of exploding offers. Haruvy et al. (2006) also study dynamics and unraveling inefficiencies of law clerk matching, using experimental and computational investigations to evaluate proposed reforms to the US system. Notably, this market is a decentralized one with no central authority designing an allocation procedure. Additionally, there is some conflict among the judges which prevents an effective coordination to improve the system. In contrast, the market for legal trainee-ships in Germany is centralized within districts of regional courts of appeals. While unraveling does not appear to happen in the allocation of lawyers in Germany, congestion is an important issue. Our paper is thus also related to common themes of the literature on markets suffering various defects (Roth and Xing, 1994; Niederle and Roth, 2003, 2009) and on how to improve the design of markets to overcome these defects (Roth and Peranson, 1999).

Two further related papers are Schummer and Vohra (2013) and Schummer and Abizada (2015). The former paper considers the assignment of landing slots to planes in the event of adverse weather. It shows the lack of incentives to report truthfully the estimated arrival times for flights under the currently used mechanism and proposes a strategy-proof alternative. That paper also highlights the restrictions that notions of incentive compatibility impose on the efficiency of the resulting mechanisms. The landing slot allocation problem as studied in those papers also differs from the lawyer allocation problem studied here. First, the paper assumes that all future arrival times are known by the airlines at the time an allocation is made. Second, the airlines have homogeneous preferences for early arrival at a single airport. So unlike in the present paper, there is only one good to be allocated in any time period.

Last, this paper is also related to other papers analyzing allocation systems in which some participants need to wait before being allocated. Braun et al. (2010) and Westkamp (2013) both study the mechanism used to allocate medical students to universities, where waiting times can be several years. However their models of the allocation procedure are static in the sense that they consider allocations for only one time period.

3 Model

This section introduces the lawyer-court many-to-one matching with waiting time problem. We abstract from complications arising from the fact that lawyers arrive sequentially over time and focus on the case in which a given set of lawyers is to be allocated to courts over several time periods. Each court can only accept a fixed number of lawyers per period.

The lawyer assignment problem consists of the following components:

1. a finite set of periods $T = \{1, ..., t_{max}\}$
2. a finite set of lawyers $I = \{i_1, ..., i_n\}$
3. a finite set of courts $C = \{c_1, ..., c_m\}$
4. a matrix of court capacities $q = (q_{c,t})_{c \in C, t \in T}$
5. lawyers' (strict, rational) preferences $P = (P_i)_{i \in I}$ over $C \times T \cup \{\emptyset\}$, with R_i denoting weak preferences of lawyer i.[26] The domain of preference profiles is denoted $\mathcal{P}$.
6. a list of courts' priority rankings, $\succ = (\succ_c)_{c \in C}$ over I.[27]

[26]This means that $(c, t)R_i(c', t')$ if and only if either $(c, t)P_i(c', t')$ or $(c, t) = (c', t')$.

[27]These can be thought of as a single score as a function of a lawyer's grade, waiting time and social factors, such as place of birth, current residence or place of study. Since we consider a static setting, we will not consider how these priority rankings might change.

We call $(T, I, C, q, P, \succ)$ an instance of a lawyer-court matching with waiting time problem. A **contract** is a triplet $x = (i, c, t) \in I \times C \times T$, specifying a lawyer, a court and the time at which the lawyer begins her trainee-ship at the court. Let $X \subseteq I \times C \times T$ be the set of all feasible contracts. For contract $x = (i, c, t)$ we denote by x_I the lawyer appearing in x, i.e. $x_I = i$. Similarly we denote by x_C and x_T the court and the time period of assignment appearing in contract x, i.e. $x_C = c$ and $x_T = t$. Further, let Y_I be the set of lawyers appearing in some set of contracts $Y \subseteq X$, that is $Y_I = \{i \in I \mid \exists y \in Y s.t.\, y_I = i\}$.

A subset of contracts $Y \subseteq X$ is an **allocation** if for all $i \in I$, $|\{y \in Y : y_I = i\}| \in \{0, 1\}$ and for all $c \in C$ and $t \in T$, $|\{y \in Y : y_C = c\}| \leq \sum_t q_{c,t}$. In words, an allocation is a set of contracts such that no lawyer appears more than once and there are not more contracts of a court for some period than number of positions available at that court overall. An allocation $Y \subseteq X$ is **feasible** if for all $c \in C$ and $t \in T$, $|\{y \in Y : y_C = c, y_T = t| \leq q_{c,t}$. Hence an allocation is feasible if each court respects its time-specific capacity constraint for each period.[28] Let $\tilde{X}$ be the set of feasible allocations. For a subset of contracts Y denote by $Y(j)$ the subset of contracts in Y involving agent $j \in I \cup C$ alternatively, if j has no contract in Y then $Y(j)$ is the empty set. Furthermore if Y is an allocation and $j \in I$, let $Y_T(j)$ be the time of start of trainee-ship according to j's contract in Y. We define $Y_C(j)$ accordingly.

A contract x is **acceptable** to lawyer i if $xP_i\emptyset$. We suppose that within the set of courts C there is a court c^G such that $q_{c^G,t} = 0$ for all $t \in T$ and where $\succ_{c^G} = \succ_G$ is the (weak) ranking induced by the lawyers' grades. Similarly we denote by c^W the empty court with a ranking induced by waiting times, $\succ_W$, and by c^S the empty court inducing a ranking by social hardship, $\succ_S$. Note that this modeling choice is not appropriate for all federal states. For example, Hamburg uses a single score to determine which lawyers are allocated.

There are two possible interpretations of our model, the myopic and the fully dynamic interpretation. Under the fully dynamic interpretation, akin to models with overlapping generations of agents, we suppose that in the initial period $t = 1$ it is already determined how many future agents there are, when they "arrive",[29] what their preferences are and how they are ranked. Over short horizons this may be a realistic possibility. However as the horizon that one considers grows, this becomes increasingly unrealistic, especially since lawyers typically only take their state exams in the period before they start applying for positions.

[28]Note that capacity used in one period does not affect capacity in future periods.

[29]A lawyer arrives in period t if all contracts involving an earlier period of allocation are unacceptable to the lawyer.

Under the myopic interpretation one only considers the problem of allocating lawyers from a single generation to courts. In that interpretation we abstract away from future generations of lawyers arriving.[30] In the myopic view the capacities of the courts beyond the current period should then not be interpreted as actual physical capacity, but as capacity which has not been reserved for future generations of lawyers. The myopic interpretation ignores the uncertainty involved in deciding how to allocate lawyers when the number and preferences of future lawyers are not yet known.

We denote by P_i not only preferences over i's assignment of a court and a time period, but also i's preferences over allocations. These preferences over allocations reflect i's preferences over assignments, so there should be no loss of clarity in this abuse of notation.

A **direct mechanism** ψ is a function $\psi : \mathcal{P} \to \tilde{X}$.[31] Hence ψ associates to each (reported) preference profile an allocation. Note that we treat courts as objects and hence they do not behave strategically, i.e. their priorities over lawyers are assumed to be given. We also take waiting time as given and do not consider changes in priorities arising from accumulated waiting time.

We now describe a few properties that lawyer preferences over the courts and the time of allocation can satisfy.

Definition. *Preferences of lawyer $i \in I$ are **weakly impatient** if for all $c \in C$, $t, t' \in T$ such that $t < t'$, then $(c,t) R_i (c,t')$.*

A lawyer's preferences are weakly impatient if a lawyer prefers to be allocated an early position at some court to a later position at the same court.

Definition. *Preferences of lawyer $i \in I$ are **strictly impatient** if for all $c, \tilde{c} \in C$ $t, t' \in T$ such that $t < t'$, then $(c,t) R_i (\tilde{c},t')$.*

Strict impatience is a strengthening of weak impatience. A lawyer having strictly impatient preferences prefers an early position at any court to a later position at *any* court. In practice we do not expect all lawyers' preferences to be strictly impatient. The reason is that there are some regions, e.g. Saxony-Anhalt, in which the average waiting time is zero, while in other regions the average waiting time is strictly positive. This would not be observed if lawyers' preferences were strictly impatient, since in that case

[30] We do however incorporate some concern for future generations by considering a basic notion of limiting harm to future generations - early filling, which we define in Subsection 3.1.

[31] In full generality the mechanism should also depend on $(T, I, C, q, \succ)$. We suppress this dependence for simplicity but will highlight whenever it becomes relevant, for example when comparing the outcome of some mechanism when a court's ranking of the lawyers has changed.

those waiting for a position in a desirable region could just switch to a less desirable region without a waiting time and thereby be better off. In addition many of the forms filled in by lawyers when applying for a position allow them to indicate a preferred entry date, which may differ from the next possible starting date.[32] It appears that many lawyers make use of the ability to postpone their starting date.[33] While delay of applicants may be for strategic reasons, delays may be rational if the lawyer plans to obtain an additional qualification, such as a one-year law degree, in the time up to the entry date. Hence it is not clear that lawyers' preferences are either weakly or strictly impatient.

3.1 Properties of Allocations and Mechanisms

To analyze the outcome of different mechanisms it is necessary to be able to talk about properties of allocations. A basic requirement of an allocation is that no lawyer should prefer the outside option to the court that she has been assigned:

Definition. *An allocation $Y \subseteq X$ is* ***individually rational*** *if for all $i \in I$, $Y(i) R_i \emptyset$. A mechanism ψ is individually rational if $\psi(P)$ is an individually rational allocation.*

Another basic requirement that any mechanism should satisfy is that it only outputs feasible allocations.

Definition. *A mechanism ψ is* ***feasible*** *if $\psi(P)$ is a feasible allocation for all $P \in \mathcal{P}$.*

We next introduce a common notion of fairness:

Definition. *An allocation $Y \subseteq X$ has* ***no justified envy****, if for any pair of contracts $x, y \in Y$ with $x_I \neq y_I$ and $(x_C, x_T) P_{y_I} (y_C, y_T)$, one of the following conditions holds: $x_I \succ_{x_C} y_I$, $x_I \succ_G y_I$, $x_I \succ_W y_I$ or $x_I \succ_S y_I$. A mechanism ψ* ***eliminates justified envy*** *if its outcome $\psi(P)$ has no justified envy for all $P \in \mathcal{P}$.*

An allocation thus is envy-free if, whenever a lawyer prefers some other lawyers' assignment, then that lawyer must have a higher priority at the court she is being assigned to than the former lawyer, a better grade, more waiting time or a higher priority based on social hardship criteria. In standard notions of fairness, usually only the court's priorities are considered. Since the policy maker in our case explicitly uses these other rankings to determine allocations, it appears natural to modify the standard notion of lack of justified envy to incorporate these additional concerns.

[32] See the application form for trainee-ships in the district of the Dusseldorf regional court of appeals, Oberlandesgericht Düsseldorf (2015a)

[33] See the weighted list of applicants for positions in Hamburg, Hanseatisches Oberlandesgericht (2015). It can be seen that many lawyers have asked their entry date to be postponed for several months.

The following definition of Pareto dominance is standard.[34]

Definition. *An allocation $Y \subseteq X$ **Pareto dominates** another allocation $\tilde{Y} \subseteq X$ if for all $i \in I$ $Y(i) R_i \tilde{Y}(i)$ and there exists at least one $i \in I$ such that $Y(i) P_i \tilde{Y}(i)$. A mechanism ψ Pareto dominates another mechanism $\tilde{\psi}$ if for all $P \in \mathcal{P}$ $\psi(P)$ Pareto dominates $\tilde{\psi}(P)$.*

It is standard to define Pareto efficiency of an allocation by the absence of another allocation that Pareto dominates it. None of the mechanisms that we study in this paper satisfy this requirement. We thus consider a weaker notion of efficiency:

Definition. *An allocation $Y \subseteq X$ is **weakly Pareto efficient** if there does not exist an individually rational allocation $\tilde{Y} \subseteq X$ such that for all $i \in I$ $\tilde{Y}(i) P_i Y(i)$. A mechanism ψ is weakly Pareto efficient if for all $P \in \mathcal{P}$ $\psi(P)$ is weakly Pareto efficient.*

As usual, a mechanism is strategy-proof if it is a dominant strategy for each agent to truthfully report her preferences to the mechanism:

Definition. *Mechanism ψ is **strategy-proof** if for all $i \in I$, for all $P \in \mathcal{P}$ and for all $\tilde{P}_i \in \mathcal{P}_i$ we have $\psi(P) R_i \psi(\tilde{P}_i, P_{-i})$. Mechanism ψ is **group strategy-proof** if, for any preference profile $P \in \mathcal{P}$, there is no $\tilde{I} \subseteq I$ and $\tilde{P}_{\tilde{I}} = (\tilde{P}_i)_{i \in \tilde{I}}$ such that for all $i \in \tilde{I}$ we have $\psi(\tilde{P}_{\tilde{I}}, P_{-\tilde{I}}) P_i \psi(P)$.*

We next define respect of improvements, first used in the matching literature by Balinski and Sönmez (1999).[35] What that property means is that a lawyer should not receive a worse assignment when her priority has increased at the courts. First we need to define what we mean by an improvement in the priority of a lawyer. In doing so, we will follow closely the presentation in Sönmez (2013).

Definition. *A priority profile $\succ$ is an **unambiguous improvement** over another priority profile $\succ'$ for lawyer i if:*

- *the ranking of i is at least as good under $\succ$ as under $\succ'$ for any court c,*

- *the ranking of i is strictly better under $\succ$ than under $\succ'$ for some court c,*

[34] Pareto efficiency is only defined with respect to the lawyers' preferences. This is justified by the fact that the courts' priorities are set administratively and therefore do not constitute real preferences. Instead they reflect a desire on by policy-makers to take into consideration grades, waiting time and social criteria. The literature on school choice similarly considers only the preferences of students for Pareto efficiency and treats schools as objects to be allocated (see Abdulkadiroğlu and Sönmez (2003)).

[35] An alternative name for respect of improvements could be priority monotonicity, since it requires that the rank of the outcome achieved by a lawyer is monotone in priority profile improvements.

- *the relative ranking of other lawyers is the same under $\succ$ and $\succ'$ for any court*

Intuitively, a priority profile improvement of some lawyer means that while all other lawyers' relative rankings among the courts are unchanged, the particular lawyer's ranking is not worse at any court (i.e. there are at most as many lawyers ranked higher than the lawyer as before) and the lawyer's ranking has improved at least at one court. Note that priority profile improvements include improvements in grades, waiting time and social hardship criteria.

Definition. *A mechanism ψ* ***respects improvements*** *if a lawyer never receives a strictly worse assignment as a result of an unambiguous improvement in her court priorities.*

Respect of improvements is a natural property to ask for. Suppose that a better grade for a lawyer leads to an unambiguous improvement in that lawyer's ranking. If respect of improvements did not hold, the lawyer would have received a less preferred position than with the worse grade. This would run counter to the view that law students should be rewarded for good performance in the exams. In addition, some may consider it to be unjust that lawyers obtain a better outcome for themselves despite having a worse grade, compared to another lawyer. Similar arguments can be made for why a mechanism should respect improvements in waiting time and social criteria.

More important, perhaps, is the implicit reliance of existing procedures on waiting time in ranking lawyers. Suppose that under some specified mechanism a lawyer improves her ranking by arriving earlier, then, if the mechanism tries to aid lawyers who arrive early by improving their ranking, this attempt to increase the welfare will hurt those lawyers if the overall mechanism does not respect improvements.

We next formalize the notion that whenever a position is not filled in some period, then no agent who would have been available that period should be assigned later. It seems reasonable to suppose that policy-makers would not be willing to allow some place at a court to go unfilled just to allow a current applicant to obtain a better allocation. This is first because lawyers provide essential work to the court at the time of their trainee-ship and second because in this way more future slots are left open which makes future lawyers (weakly) better off.[36]

[36] The notion of early filling requires that if positions are not taken in an early period, then no agent should be allocated in a later period. It thus makes sense to require early filling only if one interprets our model as involving a single cohort of students, rather than overlapping cohorts. In an extended dynamic setting one should amend the definition of early filling to allow positions to be empty even if a lawyer from a later generation takes a position at a later time. Early filling would then only rule out lawyers from the cohort appearing at a time t to take positions after that period if there are empty slots in t.

Definition. *An allocation $Y \subseteq X$ satisfies **early filling** if there is no $t \in T$ such that there exists some $c \in C$ such that $|\{y \in Y : y_T = t, y_C = c\}| < q_{c,t}$ and there exists some $i \in I$ such that $Y_T(i) > t$. A mechanism ψ satisfies early filling if for all $P \in \mathcal{P}$, $\psi(P)$ satisfies early filling.*

Early filling appears similar in flavor to the notion of no wastefulness, which is defined as:

Definition. *An allocation $Y \subseteq X$ is **wasteful** if there exists a time t, a court c and a lawyer i such that $|\{y \in Y : y_T = t, y_C = c\}| < q_{c,t}$, $Y(i) = \emptyset$ and $(c,t)P_i\emptyset$.*

Note that a Pareto efficient allocation is automatically non-wasteful. The following example show that non-wasteful and early filling are logically independent properties:

Example 1. *There are three lawyers i_1, i_2, i_3, two courts c_1, c_2 and two time periods $t = 1, 2$. Each court has a unit of capacity in each period. All contracts are acceptable to all lawyers. The allocation $\{(i_1, c_1, 2), (i_2, c_2, 2), (i_3, c_1, 1)\}$ satisfies non-wastefulness but violates early filling. The allocation $\{(i_1, \emptyset), (i_2, c_1, 1), (i_3, c_2, 1)\}$ satisfies early filling but is wasteful.*

In fact there is a fundamental conflict between early filling and non-wastefulness if it is additionally required that allocations are acceptable to the lawyers.

Lemma 1. *There is no mechanism that is individually rational, non-wasteful and satisfies early filling.*

Proof. Suppose ψ is an individually rational, non-wasteful mechanism. We show that there is an instance of a lawyer-court matching with waiting times problem in which this mechanism necessarily violates early filling. Consider the following example. There are three lawyers i_1, i_2, i_3, two courts c_1, c_2 and two time periods $t = 1, 2$. Each court has a unit of capacity in each period. The only acceptable contracts are: $\{(i_1, c_1, 2), (i_2, c_2, 2), (i_3, c_1, 1)\}$. Then $Y' = \{(i_1, c_1, 2), (i_2, c_2, 2), (i_3, c_1, 1)\}$ is the unique individually rational and non-wasteful allocation, which does not satisfy early-filling. $\square$

In the proof above, both individual rationality and non-wastefulness are required. Without individual rationality, one of the lawyers assigned in period $t = 2$ could have been assigned to an (unacceptable) earlier position. Non-wastefulness is required, since otherwise both lawyers allocated in period $t = 2$ could have been left unassigned.

Usually non-wastefulness is one of the most basic desirable properties that a matching mechanism has to possess. In our application, not being assigned in a particular federal state however likely is the result of having been accepted elsewhere. Therefore, not being assigned appears to us not to harm lawyers to an excessive extent since with a high probability they were accepted elsewhere. Failing to satisfy early-filling can however have a detrimental effect on future generations of lawyers.

4 Berlin Mechanism

We now study the procedure that is currently used in Germany to allocate lawyers to courts, mostly adopting the myopic interpretation of our model. Some aspects of that procedure are reasonably well documented, however the part describing how lawyers are allocated to courts within a period is not. While reported preferences, capacities and priorities are to be taken into account, there is no description of how these are used to find the allocation within a period. Another complication is that lawyers have many strategic options, in addition to reporting preferences over courts. For example they can decide for what entry date they wish to apply. They can refuse to accept an offer that has been made. They can report verifiable information about social status and other information that affects the priorities they will have. Because of this complexity we decide to model the procedure in a stylized manner that captures the most important features shared by the different allocation procedures as discussed and referenced in the Introduction.

The Berlin mechanism is a **two-stage procedure**. Lawyers are only able to report a ranking over the courts and fix a particular entry date to which we suppose the lawyers have applied.[37] In each given time period, the first stage of the procedure determines the set of lawyers to be considered in this time period, while in the second stage these considered lawyers are matched to open court positions.

The **first-stage lawyer selection procedure** in a given period is often detailed in the relevant regulations, as discussed earlier. This lawyer selection procedure can vary across federal states (see observations in the Introduction), nevertheless in terms of our results, these details do not matter. The important point that the lawyer selection procedure satisfies, is that it selects lawyers based solely on observable characteristics such as grades, waiting time and social criteria while ignoring preferences of the lawyers. We describe here a stylized lawyer selection procedure, which takes as input λ_G, λ_W and λ_S, which are respectively the share of positions to be assigned to lawyers based on grade, waiting time and social hardship criteria. Let $Q^t = \sum_{c=1}^{|C|} q_{c,t}$ be the total capacity of the courts in period t. For period t select the $\lfloor \lambda_S Q^t \rfloor$ lawyers ranked highest according to

[37] In Subsection 4.3 we study the question whether lawyers can strategically delay their application.

$\succ_S$.[38] Next, select the $\lfloor \lambda_G Q^t \rfloor$ lawyers ranked highest according to $\succ_G$. Finally, select the $Q^t - \lfloor \lambda_S Q^t \rfloor - \lfloor \lambda_G Q^t \rfloor$ lawyers ranked highest according to $\succ_W$.[39]

For the **second-stage allocation of lawyers to courts** within a given period we make the assumption that the lawyer-proposing deferred-acceptance algorithm of Gale and Shapley (1962) is used.[40] The DA algorithm works as follows, taking as input lawyers' reported preferences and courts' priority rankings over lawyers:

- **Step 1:** Each lawyer applies to her preferred court. Each court considers all applicants and tentatively accepts the ones it ranks highest up to its capacity. All others are rejected.

- **Step k:** Any lawyer who was rejected by a court in the previous step applies to her next most preferred acceptable court or, if all acceptable courts have already rejected her, she is assigned the outside option. Each court considers applicants it tentatively holds from the last step and those who applied in step k and tentatively accepts the ones it ranks highest up to its capacity. All others are rejected.

The DA algorithm eventually stops with all lawyers either assigned to a court or the outside option. We summarize below a number of properties of the DA mechanism that we will use throughout the paper.[41]

Theorem. *(Dubins and Freedman, 1981; Roth, 1982) The Lawyer-Proposing Deferred-Acceptance mechanism is strategy-proof.*

Theorem. *(Balinski and Sönmez, 1999) The Lawyer-Proposing Deferred-Acceptance mechanism lacks justified envy and respects improvements.*

Given the lawyer selection procedure and the DA algorithm, the **Berlin mechanism** proceeds as follows, for ascending integers $t = 1, ..., t_{max}$:[42]

[38] We define $\lfloor x \rfloor$ to be the largest integer below x.

[39] Note that $Q^t - \lfloor \lambda_S Q^t \rfloor - \lfloor \lambda_G Q^t \rfloor$ equals $\lfloor \lambda_W Q^t \rfloor$, thereby ensuring that a total of Q^t lawyers gets selected.

[40] If an unstable mechanism were used instead of the lawyer-proposing DA, then we would immediately have the result that the Berlin mechanism cannot simultaneously satisfy individual rationality, respect of improvements, non-wastefulness and lack of justified envy (see Balinski and Sönmez (1999)). Making this assumption allows us to conclude that any deficiencies we find are likely the result of the way the Berlin mechanism determines the time at which a lawyer is allocated to a court. Consequently, this can be considered as the "most conservative" assumption.

[41] Note that when discussing the elimination of justified envy under the DA mechanism we refer to the standard definition of justified envy, which does depend on the rankings over grade, waiting time and social criteria.

[42] We consider a stylized version of the Berlin mechanism. It does not allow lawyers who could not be assigned at some period to later be allocated. Instead such lawyers are assigned the outside option. We make this assumption here for simplicity. In practice, such lawyers may be considered again by the mechanism in later rounds, with the caveat that they will not gain waiting time following a rejection.

First stage

- **Step t.a:** Select up to Q^t lawyers from the set of lawyers that so far have not been selected, according to the lawyer selection procedure.

Second stage

- **Step t.b:** Selected lawyers submit preferences over courts.[43] Apply the DA algorithm using submitted preferences of the lawyers who have so far been selected in period t and on the courts' priorities. Assign each lawyer to the court assigned under this algorithm. If there are lawyers that were assigned the outside option, go to step $t.c$. If there are no lawyers that were assigned the outside option in this step, go to step $(t+1).a$ or if $t = t_{max}$, end the procedure with all those who were not yet assigned a court being assigned the outside option.

- **Step t.c:** Select as many additional lawyers from those not yet selected as there are unassigned lawyers resulting in Step $t.b$. Repeat Step $t.b$ with those lawyers additionally selected and those that were assigned to a court before.

Intuitively the Berlin mechanism tries to allocate lawyers to the earliest possible period using the DA algorithm. If there are more lawyers than seats at courts for the earliest period there is a first step that determines the set of lawyers to be allocated to courts at the earliest date based on grades, waiting time and social criteria. If at some point a lawyer is allocated to the outside option the mechanism selects an additional lawyer to be allocated in the earliest possible period. Once all positions in the earliest possible period have been filled, the same process is repeated for the subsequent period. From the description of the Berlin mechanism it follows that the allocation it produces is feasible and that no lawyer receives an allocation that is worse than the outside option.

Lemma 2. *The Berlin mechanism is individually rational and feasible.*

In addition to the previous description of the allocation procedure, there are some peculiarities that may affect its performance, which for now we abstract from in the

[43] In practice lawyers submit their ranking over courts the first time they apply for a position. However it is conceivable that lawyers might contact the *regional court of appeal* to change those submitted preferences. However if such behavior is infrequent it is sensible to assume that preferences over courts are submitted only once by the lawyers. Some federal states explicitly allow lawyers to change their ranking over courts until a position has been offered to them. Allowing lawyers to submit their ranking over courts after the time at which they are allocated has been determined simplifies the strategic analysis of the Berlin mechanism. Given that DA is strategy-proof for lawyers, they will have an incentive to report their ranking over courts in a way that is consistent with their preferences over courts and time, by reporting possibly different rankings over courts for different periods.

following theoretical discussion. First, truncated court preferences: In some federal states only two more courts in addition to the most-preferred one can be reported (sometimes with no ordering possible) and if the lawyer is not allocated to any of these three, then her preference list is randomly filled with non-listed courts.[44] Second, endogenous court priorities: lawyers can report a *verifiable* special social connection to some courts, e.g. a spouse or other relatives living in that region etc., leading to higher priority at that court. Third, refusals to accept positions: lawyers are informed of their allocated court, but they can refuse to accept that position. Refusing lawyers are replaced by those still on the waiting list. Usually, refusals lead to non-accrual of waiting time.

4.1 Deficiencies of the Berlin Mechanism

The algorithm as currently used has a number of flaws, mainly associated to the fact that t-preferences are not considered when determining which lawyers are to be allocated in a given time period. While lawyers are able to report different rankings over courts for different periods, any information concerning the trade-off between waiting and obtaining a better court is not used by the mechanism. This has important implications for the efficiency of the mechanism.

Proposition 1. *The Berlin mechanism is not weakly Pareto efficient.*

Proof. Consider the following example.

Example 2. $C = \{c_1\}$*,* $I = \{i_1, i_2\}$*,* $T = \{1, 2\}$*,* $q_{c_1,1} = q_{c_1,2} = 1$ *and* $i_1 \succ_{c_1} i_2$*. But lawyers arrive in the first period and their preferences are* $(c_1, 2)P_{i_1}(c_1, 1)$ *and* $(c_1, 1)P_{i_2}(c_1, 2)$*. Due to the two-stage procedure the higher ranked lawyer* i_1 *is considered for the first period, reports* c_1 *as acceptable and is allocated. In period two* i_2 *is selected and allocated. Then, the outcome of the Berlin mechanism is* $\{(i_1, c_1, 1), (i_2, c_1, 2)\}$*, which is strictly worse for both lawyers than* $\{(i_1, c_1, 2), (i_2, c_1, 1)\}$*.*

□

Lack of weak Pareto efficiency means that under the Berlin mechanism there could be situations under which every single lawyer could be made better off. In the above example the inefficiency stems from the fact that the exogenous lawyer selection procedure in effect determines the final allocation of lawyers to courts. Since this allocation does not depend on lawyers' preferences at all, it is not surprising that there are many lawyers that could be made better off. We have the following result.

[44]It is well known that the DA mechanism is more manipulable if lawyers report a ranking over only k courts than if lawyers report a ranking over $k' > k$ courts, see Pathak and Sönmez (2013).

Proposition 2. *The Berlin mechanism does not eliminate justified envy and does not respect improvements.*

Proof. Consider the following lawyer assignment problem.

Example 3.

[Berlin mechanism does not eliminate justified envy] There are two periods, $t = 1, 2$. *We have three lawyers* $I = \{i_1, i_2, i_3\}$ *and two courts* $C = \{c_1, c_2\}$. $q_{c_1,1} = q_{c_1,2} = q_{c_2,1} = 1$ *and* $q_{c_2,2} = 0$. *Court priorities are* $i_1 \succ_c i_2 \succ_c i_3$ *for all* $c \in C$. *Lawyer preferences are*

$$\begin{aligned} i_1 &: (c_1,1)P_{i_1}(c_1,2)P_{i_1}(c_2,1) \\ i_2 &: (c_1,1)P_{i_2}(c_1,2)P_{i_2}(c_2,1) \\ i_3 &: (c_1,1)P_{i_3}(c_2,1)P_{i_3}(c_1,2). \end{aligned}$$

In period 1, *in the first stage the two lawyers with highest priority* (i_1 *and* i_2), *regardless of their preferences, are selected to be allocated to the two open spots in the first period. Lawyer* i_3 *is put on hold, increases her waiting time, and will be reconsidered in the next period. In the second stage of period* 1 *lawyers can report their preferences considering only contracts for this time period* $t = 1$. *Based on these preferences,* i_1 *and* i_2 *are matched to their favorite courts, respecting their priority, and using the deferred acceptance mechanism. In period* 2, *there is only* i_3 *who is then allocated.*

Therefore the Berlin mechanism produces the following (unique) outcome $X^{Berlin} = \{(i_1,c_1,1),(i_2,c_2,1),(i_3,c_1,2)\}$ *for all* $c \in C$. *This outcome is not fair since there exists justified envy of* i_2, *i.e.* $(c_1,2)P_{i_2}(c_2,1)$, *although* $i_2 \succ_c i_3$.[45]

[Berlin mechanism does not respect improvements] Consider the previous set-up. If courts' priority orders are changed to $\succ'$, *s.t.* $i_1 \succ'_c i_3 \succ'_c i_2$, *then the resulting allocation under the Berlin mechanism is* $X^* = \{(i_1,c_1,1),(i_2,c_1,2),(i_3,c_2,1)\}$. *If* i_2 *improves, e.g. with a better grade, such that the old priority ranking,* $\succ$, *is recovered, then* X^{Berlin} *would result and* i_2 *would be worse off. Hence the algorithm does not respect improvements.*

□

Finally, we show that the Berlin mechanism may be wasteful.

Proposition 3. *The Berlin mechanism is wasteful.*

Proof. Consider the following lawyer assignment problem.

[45] Note that the allocation $X^* = \{(i_1,c_1,1),(i_2,c_1,2),(i_3,c_2,1)\}$ is preferred by i_2 and i_3 and weakly preferred by i_1 to X^{Berlin} and hence Pareto dominates it, despite equal courts' rankings of lawyers.

Example 4. *Suppose there are three lawyers i_1, i_2 and i_3, two courts c_1, c_2 and two time periods $t = 1, 2$. Each court has one unit of capacity in each time period. All courts rank lawyers the same: $i_1 \succ_c i_2 \succ_c i_3$. Lawyers i_1 and i_3 are strictly impatient and prefer c_1 over c_2 in both periods. Lawyer i_2 is (strictly) impatient and finds only c_1 acceptable. Then the outcome of the Berlin mechanism under truth-telling is: $Y' = \{(i_1, c_1, 1), (i_2, \emptyset), (i_3, c_2, 1)\}$. Note that because i_2 does not find $(c_1, 1)$ acceptable, he is assigned the outside option in $t = 1$. This is wasteful since court c_1 has an empty position in period $t = 2$ which lawyer i_2 prefers to being unassigned.*

□

Note that Proposition 3 is a corollary of Lemma 1 which shows that no individually rational mechanism is both non-wasteful and fills positions early and Proposition 6, which shows that the Berlin mechanism fills positions early.

4.2 Desirable Properties of the Berlin Mechanism

We have seen that the Berlin mechanism is not weakly Pareto efficient, does not eliminate justified envy, does not respect improvements and further is wasteful for general preferences. One question that could be considered is whether there exists a class of preferences for which the Berlin mechanism eliminates justified envy and respects improvements. As it turns out for preferences which are *strictly impatient* the currently used allocation procedure always delivers an allocation without justified envy and respects improvements. This is summarized in the following proposition.

Proposition 4. *Suppose the preferences of each lawyer are strictly impatient and all contracts are acceptable to all lawyers and courts. Then the Berlin mechanism eliminates justified envy.*

Proof. Suppose first lawyers are strictly impatient, all contracts are acceptable to all lawyers and courts and that the Berlin mechanism does not eliminate justified envy. Then there exist lawyers i_1, i_2, contracts x, y resulting in the Berlin mechanism with $x_I = i_1$ and $y_I = i_2$ such that $y P_i x$ and $i_1 \succ_{x_C} i_2$, $i_1 \succ_G i_2$, $i_1 \succ_W i_2$, and $i_1 \succ_S i_2$. Since all contracts are acceptable we have that $x \neq \emptyset$. Since i_1 is higher ranked than i_2 in terms of grade, waiting time and social criteria, the Berlin mechanism must yield contracts such that $y_T \geq x_T$. Strict impatience rules out that $y_T > x_T$, as otherwise i_1 would not prefer y to x. Thus we have $y_T = x_T$. But in period x_T the deferred-acceptance algorithm is used, which is known to eliminate justified envy. This contradicts $i_1 \succ_{x_C} i_2$. □

Note that we need that all lawyer find every court acceptable. If this were not the case, it could happen that lawyer i is selected for a period earlier than lawyer j, but left unassigned. This can happen if there are some courts that lawyer i finds unacceptable and if j later obtains a position at a court that is acceptable. While this result is somewhat encouraging, one should note that in practice it is not obvious that lawyers have preferences that are strictly impatient and find all courts acceptable, which we have already argued does not appear likely.

Unfortunately the above logic cannot be used to show that under strict impatience and acceptability of all contracts, the Berlin mechanism respects improvements. To see this, consider the following example:

Example 5. *$I = \{i_1, i_2, i_3, i_4\}$, $C = \{c_1, c_2\}$, $t = 1, 2$ and $q_{c_1,1} = q_{c_2,1} = q_{c_1,2} = q_{c_2,2} = 1$. Preferences of the lawyers are given by $(c_1, 1)P_i(c_2, 1)P_i(c_1, 2)P_i(c_2, 2)$ for all $i \in I$. There are two priority profiles, $\succ$ and $\tilde{\succ}$, given by:*

$$\begin{array}{lcl} \succ_{c_1}: i_3, i_2, i_1, i_4 & & \tilde{\succ}_{c_1}: i_3, i_2, i_1, i_4 \\ \succ_{c_2}: i_1, i_2, i_3, i_4 & \textit{and} & \tilde{\succ}_{c_2}: i_1, i_2, i_3, i_4 \\ \succ_{G}: i_1, i_2, i_3, i_4 & & \tilde{\succ}_{G}: i_2, i_1, i_3, i_4 \\ \succ_{W}: i_2, i_3, i_4, i_1 & & \tilde{\succ}_{W}: i_2, i_3, i_4, i_1 \end{array}$$

Note that the priorities only differ in that i_2 has a better grade under $\tilde{\succ}$. We consider a lawyer selection procedure with $\lambda_G = 0.5$ and $\lambda_W = 0.5$. As $\lambda_S = 0$ by implication, the lawyer selection procedure first selects the lawyer ranked highest according to grade and the highest ranked remaining lawyer according to waiting time. Hence under $\succ$ lawyers i_1 and i_2 are selected and lawyer i_2 gets $(c_1, 1)$. Under $\tilde{\succ}$ lawyers i_2 and i_3 are selected. So i_2 gets $(c_2, 1)$ under $\tilde{\succ}$, meaning that the improvement in the ranking of lawyer i_2 has made her worse off, despite the fact that all agents have strictly impatient preferences.

For the Berlin mechanism to respect improvements we need a further assumption: namely that a single ranking determines which lawyers are allocated for each period.

Proposition 5. *Suppose the preferences of each lawyer are strictly impatient, all contracts are acceptable to all lawyers and courts and $\lambda_G = 1$. Then the Berlin mechanism respects improvements.*

Proof. Let $\tilde{\succ}$ be an unambiguous improvement over $\succ$ for lawyer i and let $x, \tilde{x}$ be the respective assignments obtained under the Berlin mechanism. For a contradiction suppose $xP_i\tilde{x}$. There are three cases. First, suppose $x_T > \tilde{x}_T$. By the Berlin mechanism and all contracts being acceptable there is a constant number $Q_t = \sum_c q_{c,t}$ of agents allocated in period t, which are the $\sum_{s=1}^{t-1} Q_s + 1$ to $\sum_{s=1}^{t} Q_s$ highest ranked agents according to either $\succ_G$ or $\tilde{\succ}_G$. Since $\tilde{\succ}$ is an unambiguous improvement, we must have $\tilde{x}_T \geq x_T$, a contradiction. Second, suppose $\tilde{x} = \emptyset$. From $xP_i\tilde{x}$ it follows that $x \neq \emptyset$. But since all contracts are acceptable, i under $\tilde{\succ}_G$ cannot be ranked higher than $\sum_{s=1}^{t_{max}} Q_s$. But then it must be ranked even lower under $\succ_G$ implying that $x = \emptyset$, a contradiction. Third, suppose $x_T = \tilde{x}_T$. But the deferred-acceptance algorithm satisfies respect of improvements, which contradicts $xP_i\tilde{x}$. □

The Berlin mechanism in Step 1a selects Q_1 lawyers to be allocated via the deferred acceptance algorithm to positions in period $t = 1$. For each unfilled position, another lawyer is selected. Hence either the position will be filled and the algorithm moves to the next period or all remaining lawyers consider the unfilled position to be unacceptable. In the former case the algorithm fills all position for period $t = 1$. In the latter case it does not fill all positions in period $t = 1$ but all lawyers have been assigned to either a position or the outside option. Hence in that case no lawyer will be allocated to a later period. This argument can be extended to any subsequent period, so that either all positions for that period are filled or no positions in subsequent periods are filled. As a result, the final assignment obtained by the Berlin mechanism fills positions early. We summarize this finding in the following proposition.

Proposition 6. *The Berlin mechanism fills positions early.*

One of the deficiencies of the Berlin mechanism was that it sometimes wasted positions. If we require all lawyers to find all positions acceptable, then this is no longer the case.

Proposition 7. *If all agents find all contracts acceptable, then the Berlin mechanisms is non-wasteful.*

Proof. Let Y' be the outcome of the Berlin mechanism for some lawyer-court matching with waiting time problem. Suppose that some lawyer i is not assigned under the Berlin mechanism, but that there exists c, t such that $|\{y \in Y' | y_C = c, y_T = t\}| < q_{c,t}$. By assumption we have that $(c,t)P_i\emptyset$. Note that i cannot have been selected at a step $t' \geq t$, since the fact that there was an empty position at court c for time t implies that more lawyers would have been selected until all positions in period t were filled. In particular, i would have been selected eventually. But then, since i finds (c,t) acceptable, i would have been assigned to it in that step. Hence i must have been selected earlier. Furthermore, i cannot have been selected in step $t' < t$. If i had been selected, i would have been assigned since i finds all courts acceptable. Hence we have a contradiction. □

4.3 Strategic Delay under the Berlin Mechanism

So far we have mainly adopted the myopic interpretation: there is a single cohort of lawyers who simultaneously apply for positions for their legal trainee-ship. This effectively assumes that all lawyers need to apply at the same time. However the Berlin mechanism may lead to incentives for strategically delaying an application. For simplicity, we again abstract away from future generations of lawyers, but allow lawyers to choose the time at which they submit their application. Note that the Berlin mechanism, by its nature, can accommodate agents submitting preferences at various points in time. We adapt the Berlin mechanism by inserting at the very beginning of the Berlin mechanism a step 0, in which each lawyer reports a desired starting time $\tau_i \in T$.

The difference to before is that in each period only those lawyers who wished to be allocated before or in that period are considered in the lawyer selection procedure. Under the Berlin mechanism with reports of starting time, the strategy of each lawyer is now a starting time τ_i as well as her preferences over courts for each period.

The following example shows that agents may have an incentive to delay submitting their preferences:

Proposition 8. *Under the Berlin mechanism, agents have incentives for delaying their application.*

Proof. Consider the following example.

Example 6. *There are two periods, $t = 1, 2$. We have lawyers $I = \{i_1, i_2, i_3\}$. There are two courts, i.e. $C = \{c_1, c_2\}$. $q_{c_1,1} = q_{c_1,2} = q_{c_2,1} = 1$ and $q_{c_2,2} = 0$. Court (as well as grade, waiting time and social) priorities are $i_1 \succ_c i_2 \succ_c i_3$ for all $c \in C$. Lawyer preferences are*

$$\begin{aligned} i_1 &: \quad (c_1,1)P_{i_1}(c_1,2)P_{i_1}(c_2,1) \\ i_2 &: \quad (c_1,1)P_{i_2}(c_1,2)P_{i_2}(c_2,1) \\ i_3 &: \quad (c_1,1)P_{i_3}(c_2,1)P_{i_3}(c_1,2) \end{aligned}$$

If all lawyers submit their desired starting time $\tau_i = 1$, the resulting allocation is $\{(i_1, c_1, 1), (i_2, c_2, 1), (i_3, c_1, 2)\}$. However if lawyer i_2 instead reports $\tau_2 = 2$, the outcome of the Berlin mechanism is $\{(i_1, c_1, 1), (i_2, c_1, 2), (i_3, c_2, 1)\}$, which is preferred by lawyer i_2 to the outcome from applying in period $t = 1$. Therefore, lawyer i_2 has an incentive to delay her application.

□

In practice, incentives for strategic delay may be muted by the (uncertain) arrival of future generations of lawyers. If there are sufficiently many highly ranked future generations of lawyers arriving in period $t = 2$, then by delaying her application, agent i_2 might not be assigned at all or later. The motivation of delaying the application in this example is for strategic reasons: it allows lawyer i_2 to obtain a more preferred allocation. In practice students might also wish to delay their entry date for non-strategic reasons. This could happen when they wish to do a PhD or a masters degree before starting their trainee-ship. In such cases the lawyers would have a preference of starting late.

Allowing lawyers to choose the time period in which they apply may alleviate some concerns regarding the negative properties of the Berlin mechanism. However the following example shows that there are equilibria under the Berlin mechanism that are not weakly Pareto efficient.

Proposition 9. *There are Nash equilibrium outcomes under the Berlin mechanism with strategic delay that are weakly Pareto inefficient.*

Proof. Consider the following example.

Example 7. $C = \{c_1, c_2\}$*,* $I = \{i_1, i_2\}$*,* $q_{c_1,1} = q_{c_2,2} = 1$*,* $q_{c_1,2} = q_{c_2,1} = 0$ *and* $i_1 \succ_{c_1} i_2$*,* $i_2 \succ_{c_2} i_1$ *and* $i_1 \succ_G i_2$ *and* $\lambda_G = 1$*. Preferences are:* $(c_2, 2)P_{i_1}(c_1, 1)$ *and* $(c_1, 1)P_{i_2}(c_2, 2)$*. Let* $\tau_1, \tau_2 \in \{1, 2\}$ *be the desired starting dates of the two lawyers, respectively. Note that reported preferences over courts are not relevant in this example. Then* $\{(\tau_1 = 1), (\tau_2 = 2)\}$ *is a Nash equilibrium strategy profile. The outcome associated with this strategy profile is* $\{(i_1, c_1, 1), (i_2, c_2, 2)\}$*. To see that this strategy profile is indeed a Nash equilibrium, suppose* i_1 *deviated to report* $\tau_1 = 2$*. Then no lawyer would be allocated in the first period. In the second period, lawyer* i_2 *would still be allocated to* c_2 *due to her higher priority at the court. Lawyer* i_1 *would be left unallocated. Hence* i_1 *does not gain from this deviation. Next suppose* i_2 *deviates to report* $\tau_2 = 1$*. Then only* i_1 *is selected to be allocated in the first period, while* i_2 *is still allocated in the second period. Hence* i_2 *is indifferent. Thus* $\{(t_1 = 1), (t_2 = 2)\}$ *constitutes a Nash equilibrium. To see that this is not Pareto efficient, note that if* i_1 *and* i_2 *switched allocations such that* $\{(i_1, c_2, 2), (i_2, c_1, 1)\}$*, both would be better off.*

□

Note however that there are multiple equilibria in the example we considered. For example the profile $\{(\tau_1 = 2), (\tau_2 = 1)\}$ would result in a Pareto efficient Nash equilibrium outcome in the example used above.

5 Stable Mechanisms

5.1 Choice Functions and their Properties

In the previous section we have seen that the currently employed procedure of allocating lawyers to their trainee-ships has some serious deficiencies. In this section we propose a procedure which overcomes these problems. Our approach is to first take the court (or grade, waiting time and social) priorities as used in the current procedure and then to construct choice functions, as in the matching with contracts literature. Having constructed the choice functions we can then use the cumulative offer process of Hatfield and Milgrom (2005) to find a stable allocation. Specifying appropriate choice functions for the lawyers does not present a difficulty since a lawyer will simply choose her most preferred contract from the set of available contracts. The choice functions for the courts are somewhat harder to define.

We will denote general choice functions of some agent $j \in I \cup C$ as Ch_j which associates for each offer set $Y \subseteq X$ some contracts involving j. When we write $\mathrm{Ch}_i(Y)$ then the choice function of an agent $i \in I$ from the offer set Y is meant, whereas $\mathrm{Ch}_c(Y)$ denotes the choice function of a court $c \in C$ from the offer set. A lawyer i's choice function $\mathrm{Ch}_i(Y)$ specifies for each set of contracts $Y \subseteq X$ which contract the lawyer chooses and is given by

$$\mathrm{Ch}_i(Y) \equiv \max_{P_i} Y.$$

The above formulation says that lawyer i will choose from set Y the contract naming lawyer i that is maximal according to the lawyer's preferences P_i. If Y does not contain a contract with i then $\mathrm{Ch}_i(Y) = \emptyset$.

While there are many possible choice functions that are conceivable for the courts, we restrict attention to slot-specific choice functions as in Kominers and Sönmez (2016). Each court c has a set S_c of slots where $|S_c| = \sum_{t \in T} q_{c,t}$. Each slot $s \in S_c$ has an associated priority ordering Π_c^s over the set of contracts involving court c, where we denote the profile of slot-specific priority orderings of court c by $\Pi_c = \cup_{s \in S_c} \Pi_c^s$. In our setting it is natural to suppose that each court has $q_{c,t}$ slots of type t. We let S_c^t be the set of slots of type t and thus we have $S_c = \cup_{t \in T} S_c^t$. Furthermore for each court c there is a precedence order $\rhd_c$ over slots in S_c. The interpretation of $\rhd_c$ is that for slots $s, s' \in S_c$ if $s \rhd_c s'$ then slot s is filled before slot s', where we make precise what filling a slot before another one means below. Given the slot-specific priorities and the precedence order over slots, a court's slot-specific choice function $\mathrm{Ch}_c(Y; \rhd_c, \Pi_c)$ is constructed as follows. Consider slots in order of their precedence $\rhd_c$. Each slot s chooses its most preferred contract according to Π_c^s from those contracts that have been offered and are not yet associated to any lawyer chosen by any slot with higher precedence.

For court c the model set-up does not prescribe a unique slot-specific choice function that is consistent with the priority $\succ_c$ and the time-specific capacity constraints. While we know a court's priority ordering over lawyers and its capacity constraints $q_{c,t}$, this does not imply a single slot-specific choice function. There are potentially many different slot-specific choice functions, differing both in the precedence order $\rhd_c$ as well as in the slot priority orders Π_s. We introduce below the **time-specific choice function** $\mathrm{Ch}_c^{ts}(\cdot) = \mathrm{Ch}_c(\cdot; \rhd^{ts}, \Pi(\succ_c))$, for which each slot of type t finds only contracts involving period t acceptable and ranks acceptable contracts according to the court's priority ordering $\succ_c$.[46] The precedence order $\rhd^{ts}$ is such that any slot of type t has precedence over any slot of type t' if $t < t'$, i.e. for all $s \in S_c^t$ and $s' \in S_c^{t'}$ such that $t < t'$ we have $s \rhd^{ts} s'$. Slots of the same type can be ordered arbitrarily without loss of generality since their priority orderings are identical. The reason for referring to this as the time-specific choice function is that it makes choices of contracts based on constraints, which specify for each time period the number of contracts that can be held. For any set of available contracts Y the choice of court c from Y, $\mathrm{Ch}_c^{ts}(Y)$, is thus given by the following procedure:

- **Step 0:** Reject all contracts $y \in Y$ with $y_C \neq c$.
- **Step $\mathbf{t} \in \{\mathbf{1}, ..., \mathbf{t_{max}}\}$:** Consider contracts $y \in Y$ with $y_T = t$. Accept one by one contracts of the highest priority lawyers according to $\succ_c$ until $q_{c,t}$ contracts have been accepted. If a contract of lawyer y_I has been accepted, reject all other contracts y' with $y'_I = y_I$. Once $q_{c,t}$ contracts have been accepted, reject all other contracts y with $y_T = t$. If there are no contracts which have not yet been considered, end the algorithm. Unless $t = t_{max}$ move to the next step $t + 1$. If $t = t_{max}$ end the algorithm.

We will make use of the following definitions of unilateral and bilateral substitutes from Hatfield and Kojima (2010):

Definition. *Contracts are **unilateral substitutes** for court c if there do not exist contracts $x, z \in X$ and a set of contracts $Y \subseteq X$ such that $z_I \notin Y_I$, $z \notin \mathrm{Ch}_c(Y \cup \{z\})$ and $z \in \mathrm{Ch}_c(Y \cup \{x, z\})$.*

[46] Since each lawyer has only one contract available for each period, this completely determines the slot's priority ordering.

Consider a situation in which for some lawyer i there is only one contract, say z, in the available set of a court that is not chosen by the court. Then the choice function of the court satisfies unilateral substitutes if and only if that contract is also not chosen when some other contract, say x, is added to the available set.

Definition. *Contracts are **bilateral substitutes** for court c if there do not exist contracts $x, z \in X$ and a set of contracts $Y \subseteq X$ such that $z_I, x_I \notin Y_I$, $z \notin \mathrm{Ch}_c(Y \cup \{z\})$ and $z \in \mathrm{Ch}_c(Y \cup \{x, z\})$.*

Bilateral substitutes is a less strict requirement on choice functions. Consider a situation in which for some lawyer i there is only one contract, z, in the available set, that is not chosen by the court. Then consider adding another contract, x, to the available set, such that the lawyer of that new contract did not previously have a contract in the available set. The court's choice function satisfies bilateral substitutes if and only if the contract z of lawyer i is still rejected out of the larger set of available contracts.

The following irrelevance of rejected contracts property as defined by Aygün and Sönmez (2012) will be needed:

Definition. *Choice functions satisfy **irrelevance of rejected contracts** (IRC) for court c if for all $Y \subset X$ and for all $z \in X \setminus Y$, we have $z \notin \mathrm{Ch}_c(Y \cup \{z\})$ implies $\mathrm{Ch}_c(Y) = \mathrm{Ch}_c(Y \cup \{z\})$.*

Irrelevance of rejected contracts simply means that the availability of contracts which are not chosen does not matter for choices.

Although we will rely on the results of Kominers and Sönmez (2016) to establish strategy-proofness of the cumulative offer process for a particular choice function, other choice functions that we introduce in this paper satisfy the law of aggregate demand, first introduced by Hatfield and Milgrom (2005):

Definition. *The choice function of court $c \in C$ satisfies the **law of aggregate demand** if for all $X' \subseteq X'' \subseteq X$, $|\mathrm{Ch}_c(X')| \leq |\mathrm{Ch}_c(X'')|$.*

The law of aggregate demand intuitively says that when more contracts are available to a court, then the court does not choose to accept fewer contracts. We can now state Lemma 3:

Lemma 3. *(Kominers and Sönmez, 2016) The time-specific choice functions satisfy bilateral substitutes and IRC.*

In general, Kominers and Sönmez (2016) have shown that slot-specific choice functions satisfy neither unilateral substitutes nor the law of aggregate demand. However since we consider a particular slot-specific choice function it could potentially satisfy these conditions. However the next two examples show that this is not the case.

Example 8. *Let* $T = \{1, 2\}$, $Y = \{(i_2, c, 2)\}$ *and* $x = (i_2, c, 1)$, $z = (i_1, c, 2)$. *Furthermore let* $i_2 \succ_c i_1$ *and* $q_{c,1} = q_{c,2} = 1$. *Then we have under a time-specific choice function* $z \notin \mathrm{Ch}_c^{ts}(Y \cup \{z\}) = \{(i_2, c, 2)\}$. *However we have* $z \in \mathrm{Ch}_c^{ts}(Y \cup \{x, z\}) = \{(i_2, c, 1), (i_1, c, 2)\}$, *which contradicts unilateral substitutes.*

Example 9. *Let* $Y = \{(i_1, c, 1), (i_2, c, 2)\}$, $i_2 \succ_c i_1$ *and* $q_{c,1} = q_{c,2} = 1$. *Then we have* $\mathrm{Ch}_c^{ts}(Y) = \{(i_1, c, 1), (i_2, c, 2)\})$ *but we also have* $\mathrm{Ch}_c^{ts}(Y \cup \{(i_2, c, 1)\}) = \{(i_2, c, 1)\}$. *Hence adding the contract* $(i_2, c, 1)$ *to the set of contracts* Y *reduces the total number of contracts chosen.*[47]

The unilateral substitutes as well as the law of aggregate demand condition is used by Hatfield and Kojima (2010) and Aygün and Sönmez (2012) to prove (group) strategy-proofness and the rural hospitals theorem for the cumulative offer process. The unilateral substitutes condition is also used to show the existence of a doctor-optimal stable matching. Nevertheless we are able to show that despite of the failure of the unilateral substitutes condition, this result continues to hold in our model. The key to this result is to assume that the preferences of lawyers satisfy the weak impatience property. With that property a situation such as the one in the example above cannot arise. There we had that a contract of lawyer i_2 for a late period was available without contracts of the same lawyer for all earlier time periods being available. Adding one of these earlier time periods then caused lawyer i_1 to be accepted when i_1 was previously rejected. If lawyers however propose early contracts before later ones, such a situation cannot arise in the cumulative offer process.

To discuss elimination of justified envy, we follow Sönmez (2013) in defining fairness of a choice function.[48]

Definition. *For any court* c, *choice function* Ch_c *is* ***fair*** *if for any set of contracts* $Y \subseteq X$, *and any pair of contracts* $x, y \in Y$ *with* $x_C = y_C = c$, $y_I \succ_c x_I$, $y_T = x_T$ *and* $x \in \mathrm{Ch}_c(Y)$, *then there exists* $z \in \mathrm{Ch}_c(Y)$ *such that* $z_I = y_I$.

In words, a choice functions of a court is fair if it chooses one lawyer's contract but not another lawyer's contract, although the latter enjoys a higher priority at that court,

[47] We thank Christian Basteck for this example and for correcting a previously incorrect lemma.
[48] Note that this is a different concept from fairness of an allocation.

this can only be if the latter lawyer has another contract which is chosen by that court. We then have the following Lemma 4:

Lemma 4. *The time-specific choice function* Ch_c^{ts} *is fair.*

We now define stability, the central concept of the two-sided matching literature since Gale and Shapley (1962).

Definition. *An allocation* $Y \subseteq \tilde{X}$ *is* ***stable*** *with respect to choice functions* $(\mathrm{Ch}_c)_{c=1}^{|C|}$ *if we have:*

1. individual rationality: $\mathrm{Ch}_i(Y) = Y(i)$ for all $i \in I$ and $\mathrm{Ch}_c = Y(c)$ for all $c \in C$; and
2. there is no court $c \in C$ and a blocking set $Y' \neq \mathrm{Ch}_c(Y)$ such that $Y' = \mathrm{Ch}_c(Y \cup Y')$ and $Y' R_i Y$ for all $i \in Y'_I$.

Hence an allocation is stable if each lawyer prefers the assignment to being allocated no contract, each court chooses its assignment over some subset of that assignment and there is no set of contracts such that a court would rather choose that set of contract, the blocking set, when this and the allocation are available, such that the lawyers having contracts in the blocking set weakly prefer those contracts over their assignment. Under the assumption that courts use the time-specific choice function $\mathrm{Ch}_c^{ts}(\cdot)$ stable allocations are feasible. Stability is not a desiderata per se in our model. In the original literature on two-sided matchings stability was seen as important in explaining whether matching procedures would systematically lead to unraveling (Roth, 1984, 1991). In our case the regional courts are not strategic players and the priorities according to which they evaluate lawyers are determined by the mechanism designer. This precludes the possibility of courts contracting with lawyers around the centralized mechanism. However, stability matters in our context as stability implies other desirable properties of mechanisms. An allocation $Y \subseteq \tilde{X}$ is the **lawyer-optimal stable allocation** if every lawyer weakly prefers it to any other stable allocation.

5.2 Cumulative Offer Process

We now introduce the **cumulative offer process** (COP) as defined in Hatfield and Kojima (2010), which is a generalization of the deferred-acceptance algorithm of Gale and Shapley (1962).

The cumulative offer process takes as input the (reported) preferences of the lawyers as well as the choice function of each court.

- **Step 1:** One (arbitrarily chosen) lawyer offers her first choice contract x_1. The court that is offered the contract, $c_1 = (x_1)_C$, holds the contract if it is acceptable and rejects it otherwise. Let $A_{c_1}(1) = \{x_1\}$, and $A_c(1) = \emptyset$ for all $c \neq c_1$.

In general,

- **Step k $\geq$2:** One of the lawyers for whom no contract is currently held by any court offers her most preferred contract, say x_k, that has not been rejected in previous steps. Let $c_k = (x_k)_C$, hold $\mathrm{Ch}_c\left(A_{c_k}(k-1) \cup \{x_k\}\right)$ and reject all other contracts. Let $A_{c_k}(k) = A_{c_k}(k-1) \cup \{x_k\}$ and $A_c(k) = A_c(k-1)$ for all $c \neq c_k$.

Now we apply Theorem 1 of Hatfield and Kojima (2010) to show that the cumulative offer process, as just described, in conjunction with the time-specific choice function produces a stable allocation.

Theorem. *[Hatfield and Kojima (2010)] Suppose the choice functions of the court used in the cumulative offer process satisfy bilateral substitutes. Then the cumulative offer process produces a stable allocation.*

The existence of a stable matching is the minimum requirement that we ask of an algorithm. By the above result and the fact that the time-specific choice functions satisfy bilateral substitutes, using the time-specific choice functions when running the COP yields a stable allocation. Hatfield and Kojima (2010) further show that if one strengthens the assumptions to unilateral substitutes for the choice functions used, then one can show that the cumulative offer process produces the lawyer-optimal stable allocation. In our case however the time-specific choice functions do not satisfy unilateral substitutes.

Nevertheless one can adapt Theorem 4 of Hatfield and Kojima (2010), as modified by Aygün and Sönmez (2012), which is used in Theorem 5 of Hatfield and Kojima (2010) to show the existence of a lawyer-optimal stable allocation (doctor-optimal in their terminology). To do so, it is sufficient to make an assumption on the preferences of the lawyers, rather than on the choice functions used by the courts. Namely we will assume that lawyers are weakly impatient. Previous results in the matching with contracts literature usually proceeded by restricting the choice functions used by the side of the market which could accept multiple contracts to obtain results, while placing essentially no restrictions on the other side of the market. Here we depart from this approach and relax the restrictions placed on the choice functions used by the side of the market which can accept several contracts (the courts) and instead put some restrictions on the single-contract side (lawyers) of the market. Both approaches, as we will see, lead to similar results.

Lemma 5. *A contract z that is rejected by a court c at any step of the cumulative offer process using the time-specific choice function Ch_c^{ts}, cannot be held by court c in any subsequent step.*

The key to our proof of this result lies in the specific choice function that we use. This causes lawyers, when a contract of theirs is rejected, to either propose to a new court or to propose to some court at which the lawyer was previously rejected. So if some court c has multiple offers, say z and z' of some lawyer i and holds z, then it will, when receiving a new contract offer from some other lawyer j, never reject z while simultaneously accepting z'. In the proof we heavily rely on Aygün and Sönmez (2012).

With this result in hand, we can now state the following lemma:

Lemma 6. *Suppose lawyer preferences are weakly impatient. The outcome of the cumulative offer process using the time-specific choice function Ch_c^{ts} produces the lawyer-optimal stable allocation.*

The proof is essentially the same proof as the one of the corresponding Theorem 5 in Hatfield and Kojima (2010) and Aygün and Sönmez (2012). Assuming weak impatience again allows us to relax the unilateral substitutes assumption and instead use the time-specific choice functions which only satisfy the bilateral substitutes assumption. The reason that this works is that because of weak impatience, any sets of available contracts that the courts will have to make choices from are sets such that if a contract x of some lawyer i is available for period t, then contracts for any earlier and feasible period for that lawyer i will also be available. On this restricted domain of sets of available contracts unilateral substitutes essentially holds for the time-specific choice function, allowing the proofs by Hatfield and Kojima (2010) to go through, with some modifications. Note that we only needed to make use of the assumption of weak impatience for proving Lemma 6.

The result in Lemma 6 is a new result, which is not implied by any of the results in Kominers and Sönmez (2016), since they consider more general slot-specific choice functions than we do here. For general slot-specific choice functions a lawyer-optimal stable allocation is not guaranteed to exist and even when such an allocation exists, the COP is not guaranteed to find it. Lemma 6 above shows that under weak impatience, a lawyer-optimal stable allocation is guaranteed to exist and that it is found by the COP. The following example shows that without weak impatience, the existence of a lawyer-optimal stable allocation is no longer guaranteed.

Example 10. *Let i_1 and i_2 prefer $(c, 2)$ to $(c, 1)$ and assume $i_1 \succ i_2$ with $q_{c,1} = q_{c,2} = 1$. Then the allocation $Y = \{(i_1, c, 1), (i_2, c, 2)\}$ is stable, while the COP produced the also* *stable allocation* $Y' = \{(i_1, c, 2), (i_2, c, 1)\}$. *Notice that* i_2 *prefers* Y, *while* i_1 *prefers* Y', *i.e. neither allocation is weakly Pareto efficient.*

5.3 Properties of the Time-specific Lawyer Offering Stable Mechanism

The time-specific lawyer offering stable mechanism (TSLOSM), ψ^{ts}, is defined to be that mechanism which associates with each preference profile the outcome of the COP using the time-specific choice functions. We will refer to this mechanism as the time-specific stable mechanism. We have the following result:

Proposition 10. *The time-specific lawyer offering stable mechanism is stable and (group) strategy-proof. If in addition lawyers' preferences are weakly impatient, then the time-specific stable mechanism is lawyer-optimal stable.*

Stability and (group) strategy-proofness follow directly from Theorem 3 in Kominers and Sönmez (2016). The second part follows from applying our Lemma 6. Instead of applying Theorem 3 in Kominers and Sönmez (2016), an alternative way of obtaining the first part of the above results when lawyers' preferences are weakly impatient is to adapt results in Hatfield and Kojima (2009) making use of the fact that under weak impatience, a lawyer-optimal stable allocation is guaranteed to exist. An important corollary of the time-specific stable mechanism being group strategy-proof is that it leads to weak Pareto efficiency as in Hatfield and Kojima (2009).

Corollary 1. *The time-specific lawyer offering stable mechanism is weakly Pareto efficient.*

Proof. Suppose otherwise. Hence there exists an allocation $Y \subseteq X$ such that for all $i \in I$ we have $Y(i) P_I \psi^{ts}(P)$. Let $\tilde{P}_i$ be the preference profile for each lawyer $i \in I$ that lists $Y(i)$ as the only acceptable contract. Then we have that $\psi^{ts}(\tilde{P}) = Y$. This implies that we have found a coalition of lawyers, namely all lawyers, that can jointly deviate to make all its member strictly better off. This contradicts ψ^{ts} being group strategy-proof. □

One of the problems in the current procedure, the Berlin mechanism, is that lawyers may be worse off by improving their ranking, for example by obtaining a better grade or having waited longer. The next proposition shows that this is not the case for the cumulative offer process using the time-specific choice function.

Proposition 11. *The time-specific lawyer offering stable mechanism respects improvements.*

The intuition behind the proof of this result, which is simply an application of Theorem 4 in Kominers and Sönmez (2016), is as follows. Let $\succ_1$ be an unambiguous improvement over $\succ_2$ for lawyer i and let ψ_1 be the associated mechanism. Similarly for $\succ_2$. Suppose the COP were run initially excluding lawyer i under $\succ_1$, which will lead to some allocation X^1. After this, lawyer i proposes contracts in order of preference. This process will terminate for some contract offer x^k, which is i's assignment under the mechanism ψ_1. Running the algorithm under $\succ_2$ without lawyer i will lead to the same initial allocation X^1 since only the ranking of lawyer i has changed. Letting i propose contracts however will lead to the same rejections occurring since $\succ_1$ is an unambiguous improvement over $\succ_2$ until x^k is offered by i, which by assumption is the final allocation under $\succ_1$ but which may nevertheless be rejected under $\succ_2$. From this it follows that i cannot do worse under $\succ_1$ than under $\succ_2$. Note that the priorities based on grades, accumulated waiting time and social criteria do not directly enter the time-specific lawyer offering stable mechanism. Hence changes in those rankings will leave the outcome of the mechanism unchanged, which is consistent with respect of improvements.

This is an important result since it implies that targeted efforts to improve the allocation obtained by specific lawyers through an improvement of their ranking can never hurt these lawyers who those efforts are intended to help. One implication is that when the ranking depends positively on grades, then lawyers are rewarded for better grades by an improvement in their assignment.

The fact that the time-specific stable mechanism respects improvements has a further implication in our application. Lawyers, in the current system, may report to have a special social relationship to a court. For example, having children grants higher priority for regional courts in Bavaria, as discussed above. Consider now a game which first asks lawyers to report any such information. In a second stage, the priorities of each court would be adjusted to reflect those reports, in case the information lawyers have reported has been verified. In case lawyers do have special social relationship to a court, but do not report it, the choice function remains unaffected. Then we have the following result, which follows by noting that reporting this information leads to an unambiguous improvement in the priority of a lawyer at a court. Since the time-specific stable mechanism respects improvements, reporting this information, holding the strategies of everyone else fixed, cannot make a lawyer worse off, but may lead to an improvement. Hence the following corollary is obtained:

Corollary 2. *Each lawyer has an incentive to report verifiable information increasing her priority at a court under the time-specific stable mechanism.*

The above corollary shows that respect of improvements is closely linked to the incentive compatibility of revealing hard information.[49] Another desirable property that the time-specific stable mechanism satisfies is elimination of justified envy.

Proposition 12. *The time-specific stable mechanism eliminates justified envy.*

Proof. To see that the time-specific stable mechanism eliminates justified envy, let $x, y \in Y \subseteq \tilde{X}$ be two contracts obtained by the time-specific stable mechanism such that $x_I \neq y_I$ and $(x_C, x_T)P_{y_I}(y_C, y_T)$. Then since by the cumulative offer process y_I must have offered (y_I, x_C, x_T) at some step during the process, it must have been rejected. But the only way that (y_I, x_C, x_T) had been rejected while x was accepted is when $x_I \succ_{x_C} y_I$, which implies that the time-specific stable mechanism eliminates justified envy. $\square$

The time-specific mechanism proposed in this section however does not fill positions early. The reason is an inherent conflict between stability and early filling. To see this consider the following example.

Example 11. *There are two periods $t = 1, 2$ and two courts $c = c_1, c_2$, each with one position in each period. In the first period, there are two lawyers $I = \{i_1, i_2\}$, with common preferences for each $i \in I$: $(c_1, 1)P_i(c_1, 2)P_i(c_2, 1)P_i(c_2, 2)$. Both courts have priorities such that lawyer $i_1 \succ_c i_2$. The outcome of the cumulative offer process using the time-specific choice function results in the allocation $\{(i_1, c_1, 1), (i_2, c_1, 2)\}$, which leaves the position in period 1 at court c_2 unoccupied even though lawyer i_2 is given a position in period 2, thereby violating early filling.*

The example shows that (lawyer-optimal) stable outcomes might not satisfy the early filling properties. There is a trade-off between preferences of lawyers from different time periods: On the one hand, a mechanism finding a stable outcome over lawyers from all periods might not fill positions early and by this make future lawyers worse off. On the other hand, guaranteeing early filling could benefit future lawyers at the costs of earlier lawyers, but would not be stable and might violate other desirable properties, such as strategy-proofness.

[49] Aygün and Bo (2013) in their analysis of college admissions with affirmative action in Brazil study the incentives of different disadvantaged groups to disclose their status to the mechanism. For example, it is desired by the policy-makers in Brazil that students from ethnic minorities who went to a public high school are given higher priority. However such students may decide not to reveal their status as an ethnic minority. In the currently used procedure in Brazil they sometimes do not have an incentive to do so. The mechanism proposed by Aygün and Bo (2013) makes it optimal for students to reveal this information truthfully.

5.4 Flexible Choice Functions

The discussion in the previous sections assumed that each court c could only accept $q_{c,t}$ lawyers in time period t. This assumption was made because the number of positions at each court is determined by the budget of the federal state several periods into the future so that the courts cannot flexibly set their own capacity for each period. In this subsection we consider the possibility of allowing each court to flexibly determine how to allocate total capacity, which is assumed to be fixed over several periods. Hence, we no longer have a time-specific capacity constraint but instead have for each court a global constraint on the total number of lawyers that can be accepted. In other words, we relax the requirement that a mechanism produces a feasible allocation.

We continue to consider slot-specific choice functions and build on the time-specific choice function Ch_c^{ts} to develop the flexible choice function Ch_c^{flex}. As before there are $q_{c,t}$ copies of a slot of type t and the slots with a lower t have higher precedence. Under the time-specific choice function Ch_c^{ts} each slot's priority ordering was such that only periods of the associated time period were deemed acceptable and acceptable contracts were ranked according to $\succ_c$. Under Ch_c^{flex} each slot's priority ordering Π_s ranks highest the contracts of the associated time period. Next highest are ranked the contracts of period 1 (or period 2 in case we are considering slots of the period 1 type), followed by those contracts of the next highest period and so on.[50] Contracts of the same period are ranked according to $\succ_c$. As a consequence each slot, irrespective of its types, considers all contracts acceptable. Note that this allows for choices that violate the time-specific capacity constraints of the courts.

Since the flexible choice function is also a slot-specific choice function and since contracts, given a fixed period, are ranked according to $\succ_c$ by slots, the cumulative offer process using these choice functions is a strategy-proof, weakly Pareto efficient mechanism that eliminates justified envy and respects improvements. The proofs are similar to those for the TSLOSM and are omitted. We call this new mechanism the flexible lawyer offering stable mechanism (FLOSM) and denote it by ψ^{flex}.

FLOSM does not necessarily satisfy the time-specific capacity constraints of the courts, while the TSLOSM satisfies these constraints. Hence there may be allocations that can be reached by FLOSM that violate feasibility under TSLOSM but are preferred over allocations feasible under TSLOSM.

[50] The ranking of contracts not involving period t by slots of type t does not matter for our results, so long as all contracts involving period t are ranked highest.

Proposition 13. *Fix a lawyer matching problem* $(T, I, C, q, P, \succ)$. *Then we have for all* i *that* $\psi^{flex}(P)(i) R_i \psi^{ts}(P)(i)$.

Proposition 13 says that all lawyers weakly prefer the outcome of the flexible lawyer offering stable mechanism over the outcome of the time-specific lawyer offering stable mechanism. The intuition behind this result is that the time-specific choice function can be obtained from the flexible choice function via truncation strategies. Truncation strategies by one side of the market make the other side of the market weakly worse off. Hence we conclude that relaxing time-specific capacity constraints and suitably adapting the choice functions used by the courts has the potential of making lawyers better off.[51]

We interpret the flexible choice function as corresponding to cases in which a court is given greater budgetary freedom with respect to when to open trainee-ship positions. In practice there may be other reasons for having time-specific constraints that make adjustments to capacities over time difficult. For example, class room sizes could constraint how many lawyers may begin their trainee-ship in any given period of time.

6 Conclusion

While the above description of the Time-Specific Lawyer Offering Mechanism is rather theoretical, from a practical point of view it could be interpreted in two ways.

First, as a mechanism of perfect foresight, the matching for all lawyers at all courts in all time periods is finalized already in period 0. This requires grades, preferences etc. of arriving lawyers to be known already at time 0. If courts before time t are unacceptable to some lawyer i, then we can think of i arriving at time t. This interpretation could be realistic for short time horizons, however less so for longer horizons.

The second interpretation is that the mechanism is run every single period, whereby lawyers are allocated to positions now and in the future, ignoring future arrivals. Then, once a lawyer is allocated to a future position, this seat remains "reserved" for the current lawyer. In that case one needs to additionally analyze how many future positions one allows to be assigned today.

An interesting extension of our model would be to consider how our proposed mechanism behaves when it needs to be applied for each period over a number of periods. Dur and Kesten (2014) consider a problem in which a set of students is to be matched to colleges, but in which the set of colleges is partitioned. They show that when the assignment happens sequentially, it is inherently difficult to have a mechanism be non-wasteful,

[51] This could happen for example by allowing courts to transfer funding for trainee-ship positions over time in response to demand, rather than sticking to an exogenously given budget for each time period.

and strategy-proof while eliminating justified envy and respecting improvements. Such results would also apply in a dynamic version of our model in which the time-specific lawyer-optimal mechanism were applied repeatedly. A related problem, that we have ignored so far, is how to manage capacity. While we assumed that capacities were given exogenously, in a dynamic procedure with excess demand one may want to reserve some capacity at some courts to ensure that future agents are not unduly disadvantaged by earlier agents taking these positions. Future research should address this question.

Appendix

A Omitted Proofs

Proof. [**Lemma 4**] Suppose to the contrary that for court c and a set of contracts Y with elements $y, x \in Y$ such that $y_C = x_C = c$, $y_I \succ_c x_I$, $y_T = x_T$ and $x \in \mathrm{Ch}_c^{ts}(Y)$ but that there does not exist $z \in \mathrm{Ch}_c^{ts}(Y)$ with $z_I = y_I$. Note that in particular this implies that $y \notin \mathrm{Ch}_c^{ts}(Y)$. Then since such a z does not exist, it must be that in step $y_T = t$ of the procedure to construct Ch_c^{ts}, y has not yet been rejected. So in step t both x and y are still available. Now x is accepted in step t since $x \in \mathrm{Ch}_c^{ts}(Y)$ while y is rejected, since $y \notin \mathrm{Ch}_c^{ts}(Y)$. This contradicts $y_I \succ_c x_I$, since the procedure to construct the time-specific choice function would have selected the contract of the agent with the better ranking. □

Proof. [**Lemma 5**] Towards a contradiction let k' be the first step a court c holds a contract z that was previously rejected at step $k < k'$. As z is rejected at step k, it was on hold by court c at step $(k-1)$ or it was offered to court c at step k. In either case no other contract of lawyer z_I could be on hold by court c at step $(k-1)$. But then, since z is the first contract to be held after an earlier rejection, court c cannot have held another contract by lawyer z_I at step k. That is $z_I \notin \left[\mathrm{Ch}_c^{ts}(A_c(k))\right]_I$. Since z is rejected at step k, this means that for all $x \in \mathrm{Ch}_c^{ts}(A_c(k))$ with $x_T = z_T$, we must have $x_I \succ_c z_I$. Let $\left[\mathrm{Ch}_c^{ts}(A_c(k))\right](z_T)$ denote the set of such contracts in time z_T. Given the definition of Ch_c^{ts}, $z \in \mathrm{Ch}_c^{ts}(A_c(k'))$ implies that some contract $x \in \left[\mathrm{Ch}_c^{ts}(A_c(k))\right](z_T)$ can no longer have been under consideration in step t of the procedure to find the court's choice. But for that to have happened, it must be that some contract y with $y_I = x_I$ and $y_T < x_T$ has been accepted in step k'. But this cannot be since by assumption z is the first contract that was rejected and subsequently accepted and because x_I cannot have offered a contract in step k' since a contract of x_I was held by the court in period $k'-1$. Hence a contradiction. □

Proof. [**Lemma 6**] To prove the lemma, it is sufficient to show that for any stable allocation $X' \subseteq \tilde{X}$ and any contract $z \in X'$, contract z is not rejected by the cumulative offer algorithm when the time-specific choice function is used. To obtain a contradiction, suppose not. Let k be the first step where court $c = z_C$ rejects contract z, and let $Y = \mathrm{Ch}_c^{ts}(A_c(k))$. Then by IRC, $z \notin \mathrm{Ch}_c^{ts}(Y \cup \{z\})$. Then by lemma 5, $z_I \notin Y_I$. As k is the first step a contract in any stable allocation is rejected, every lawyer in Y_I weakly prefers their contract in Y to their contract in X' which is stable by assumption. We

then consider two cases:
Case 1: $z \notin \mathrm{Ch}_c^{ts}(Y \cup X')$. In this case, court c blocks allocation X' together with lawyers in Y_I, contradicting stability of X'.
Case 2: $z \in \mathrm{Ch}_c^{ts}(Y \cup X')$. But this cannot be, since for any $x \in \mathrm{Ch}_c^{ts}(Y)$ with $x_T = z_T$, we have for all $s < t$, $(x_I, c, s) \in Y$ by weak impatience.
Therefore the addition of contracts cannot result in z being chosen when both Y and X' are available due to the way the time-specific choice function is constructed. A contradiction. □

For the proof of Proposition 13 we make use of an associated lawyer-slot matching market as in Kominers and Sönmez (2016). A lawyer-slot matching market is constructed from a lawyer-court allocation problem in which the courts have slot-specific choice functions $\mathrm{Ch}_c(\cdot; \rhd_c, \Pi_c)$ as follows. The contract set X is extended to the set Z defined by $Z \equiv \{(x, s) : x \in X \textit{ and } s \in S_{x_C}\}$. Slot priorities $\tilde{\Pi}^s$over contracts in Z are derived from priorities Π_c^s over contracts in X. This means that $(x, s)\tilde{\Pi}^s(x', s)$ if and only if $x\Pi_c^s x'$. A lawyer's preferences $\tilde{P}_i$ over contracts in Z remain the same as preferences P_i over contracts in X with ties between the same contract at different slots broken according to the precedence order. This means that $(x, s)\tilde{P}_i(x', s')$ if and only if either xP_ix' or $[x = x' \textit{ and } s \rhd_{x_C} s']$.

Proof. **[Proposition 13]** For any instance of a lawyer-court allocation problem we construct the lawyer-slot matching problem as follows. By Theorem A.1 of Kominers and Sönmez (2016) the outcome of the lawyer offering stable mechanism in a lawyer-court matching with waiting time problem in which courts use slot-specific choice functions, corresponds to the outcome of the lawyer offering stable mechanism in the associated lawyer-slot matching market.

Suppose now that one slot $s \in S_c$ in the lawyer-slot market truncates from its priority ordering its lowest ranked contract, say x. If that contract was not part of the allocation under the lawyer offering stable mechanism without the truncation, then this truncation has no effect on the final allocation. If that contract was part of the allocation under the lawyer offering stable mechanism without the truncation, then the lawyer x_I applies to her next highest ranked slot, s'. The slot s' will either accept lawyer x_I's contract (in the process possibly rejecting another contract of another lawyer) or reject it. In either case there will be a finite chain of rejections of contracts. All lawyers involved in this rejection chain will receive a worse allocation than without the truncation of slot s according to their preferences over slots. There are now two possibilities. Either all lawyers find the new allocation worse only because of the tie-breaking induced by the precedence order $\rhd_c$. In that case the lawyers in the original lawyer-court matching with waiting time problem are unaffected by the truncation. If the new allocation is worse because of a change in the court and time period allocated to a lawyer, then lawyers in the original lawyer-court matching will be worse off.

A similar logic applies to any further truncation by any slot. Each truncation makes the lawyers weakly worse off. Consider now the case in which each slot's priorities have been truncated to only find contracts involving a time period corresponding to the slot's type acceptable. In that case the outcome of the lawyer-optimal stable mechanism in the lawyer-slot matching problem corresponds to the outcome of the lawyer-proposing stable mechanism under the time-specific choice functions. By the previous arguments, all lawyers weakly prefer the allocation without any truncations. $\square$

Chapter II

Privacy and Platform Competition

Based on Dimakopoulos and Sudaric (2018), *International Journal of Industrial Organization*, https://doi.org/10.1016/j.ijindorg.2018.01.003.

1 Introduction

Online platforms often do not charge monetary prices from users but monetize through an advertisement-based business model building on the collection and processing of user data. Typical examples include social networks (e.g. Facebook, LinkedIn), search engines (e.g. Bing, Google) or video platforms (e.g. Youtube, Vimeo). The role of user data in this context is ambiguous. From the platform perspective user data is an input factor which can be used to gain insights about users and improve the targeting of advertisement, resulting in a superior product for potential advertisers. This commodity attribute of data is mirrored to a lesser extent on the user side. Users typically accept some conditions to what extent personal data is collected and processed when using a platform service. In some cases the provision of personal data is necessary to make meaningful use of a platform service (e.g. social networks) while in other cases services do not require the collection of user data per se (e.g. search engines, mail providers, video platforms). In both cases the provision of data from a user perspective can be interpreted as a price the user is willing to accept in exchange for the use of the platform including the display of ads.[1] To put it in terms of platform economics, user data requirements exhibit price

[1] A study by the Pew Research Center (2014) shows that 91 percent of respondents agree that they lost control over how companies collect personal data while 55 percent state that they are willing to share some information in exchange for using a free service. The European Commission (2015), however, reports that 72 percent of internet users worry they provide too much data online. This indicates that users are aware and willing to exchange personal data for services, but the actual extent worries them.

characteristics on the one hand, and affect indirect network effects (e.g. targeting) at the same time.

This ambiguity makes it especially hard for policy makers as standard economic arguments might not be applicable. Indeed, the *European Data Protection Supervisor (EDPS)* argues that competition authorities should take privacy and data related aspects more into account (EDPS, 2014).[2] And indeed, recent cases demonstrate that competition authorities acknowledge the peculiarities of data-driven industries. Germany's *Federal Cartel Office (Bundeskartellamt, BKartA)* initiated investigations against Facebook in 2016 based on alleged abuse of market power. In particular, the BKartA investigates whether Facebook uses its dominant position in the market for social networks to expand the terms of service outlining how much data is collected and processed by the platform.[3] Therefore, we want to shed some light on the role of competition intensity in a two-sided market framework when users provide data and this data is monetized on the opposing market side.[4]

We analyze a setting of two competing ad-financed platforms in a two-sided market framework. On the user market side, platforms strategically set the required level of data provision, to which users have to agree to obtain access to the platform service. Platforms process this user data to sell improved ad targeting on the advertiser market side. While users incur disutility from providing data (privacy concerns, opportunity costs), they benefit from seeing more relevant ads. Users and advertisers are assumed to single-home.

Our model predicts that platforms will extract a distorted amount of data compared to the efficient benchmark. The distortion is induced through the one-sided monetization in a way that platforms do not perfectly balance the costs of data provision, i.e. privacy costs incurred by users, against the targeting benefits on both market sides, but put too much or too little weight on the benefit captured by the monetized market side. This distortion depends on the net effect of cross-group externalities as well as the degree of

[2] Whether competition authorities should incorporate aspects of privacy and data protection is, however, controversial. For arguments in favor we refer to Stucke and Grunes (2016), arguments against can be found e.g. in Cooper (2013).

[3] Bundeskartellamt, 'Bundeskartellamt initiates proceeding against Facebook on suspicion of having abused its market power by infringing data protection rules', Press Release, 2 March 2016, `http://www.bundeskartellamt.de/SharedDocs/Meldung/EN/Pressemitteilungen/2016/02_03_2016_Facebook.html`.

[4] Classical examples include ad-based business models where data is used to improve ad targeting or matching / recommendation platforms, where users are presented offers which become more relevant the more the platform knows about its users. For illustration purposes we stick to the example of targeted advertising and refer to the extension part of this paper for a more general consideration of cross-group externalities, i.e. also the possibility of users enjoying the presence of firm's offers.

competition intensity on both market sides. If targeting benefits are small or competition is weak, an inefficiently high level of data is collected. On the other hand, if competition is strong or targeting benefits sufficiently outweigh nuisance costs, too little data is collected. From the point of view of consumers the competitive level of data provision is always too high, suggesting that applying a consumer standard to online platforms leads to underprovision of personal data. The competitive equilibrium level of data provision, however, is monotone in the degree of competition intensity: the weaker the competition on either side of the market, the higher the equilibrium amount of data provision. This result is interesting because it does not follow the common two-sided platform logic that less elasticity on one side typically decreases the other side's price.

Our findings indicate that the inefficiency of data provision can be reduced by careful privacy regulation or competition policies on either market side. One interpretation of this result is that (competition) policy measures in these data-driven industries should take into account the effects they have on the extent of private data collection.

We also consider a variety of extensions to this setup. In the first one we depart from the assumption that platforms are restricted in their price setting on the user side, and allow for non-zero user prices. In fact, lifting the restriction leads to an efficient level of collected data, while user prices can be positive, negative (or zero). This gives rise to two interpretations. The first is a Coasian one, where establishing the missing market on the user side leads to an efficient outcome. This reflects the idea of Laudon (1996) that users should be adequately compensated for the provision of their data, while the problem of the 'data economy' lies precisely in the absence of such a market. The second interpretation is of counterfactual nature. In particular we argue that whenever the unrestricted model would yield positive (negative) user prices, the restricted model exhibits overprovision (underprovision) of user data as platforms can no longer adequately charge or compensate users for collecting data. The second extension considers different degrees of platform collusion and we conclude that the amount of collected data is excessively high under full collusion, while this is not necessarily the case under partial collusion. In the third extension we discuss the robustness of our results with respect to multi-homing and elastic total demand. Lastly, we demonstrate that our results naturally extend to settings with positive cross-group externalities (matching platforms).

The remaining paper is structured as follows. Section 2 relates our analysis to the existing literature. Section 3 introduces the model. Section 4 characterizes the efficient benchmark and competitive equilibrium outcomes, for which we present comparative statics in Section 5. Section 6 compares these outcomes and outlines policy implications.

In Section 7 we extend and discuss the baseline model. Section 8 concludes. Omitted proofs can be found in Appendix A, while supplementary analyses are delegated to (online) Appendix B.

2 Related Literature

Methodologically, our research is related to the literature on platform competition in general and on applications in media markets in particular. We consider a competitive setting with two-sided single-homing which has been analyzed by Armstrong (2006) in a more general framework and later extended in Armstrong and Wright (2007). However, both papers consider the case where platforms engage in two-sided pricing while non-monetary aspects (as e.g. user data) are not modelled. We also share a common component with the literature on media platforms in the sense that we, at least in our baseline model, consider the case of opposing indirect network effects, where advertisers like to reach many users but users dislike the presence of advertisers. This reflects the idea of 'peace and quiet' privacy in Posner (1981) and is a common assumption in the media literature (see Anderson and Gabszewicz (2006) for a review). This setup is used e.g. to study competition in TV markets (see e.g. Anderson and Coate (2005) or Peitz and Valletti (2008)) where platforms do not engage in targeted advertising and therefore the expected revenue per user as well as perceived nuisance are constant. Our research differs in the sense that we endogenize those indirect network effects as we let them to be affected by the level of data collected. The concept of endogenous network effects is captured in Reisinger (2012) where users spend time using platform services and platforms translate this activity into better targeting and reduced nuisance. A similar setup is presented in Bourreau et al. (2017), however the research question differs substantially. The key difference is that in our model the level of data provision is a strategic decision of the competing platforms, while in the two previously mentioned papers consumers voluntarily spend time/provide data on the platforms, which changes the competitive dynamics significantly.

We also contribute to the broader literature on efficient provision of personal data and the role of privacy as a competition instrument. The aspect of data provision being a strategic choice made by platforms is captured to some extent by Spiegel (2013) who compares commercial software (full privacy) to adware (positive privacy costs) and shows that adware is welfare superior. De Corniere and De Nijs (2016) consider a setting where a monopolistic platform auctions off advertising slots and decides whether to disclose consumer information (no privacy) or not (privacy). They show that platforms might prefer information disclosure, which comes at the cost of some consumers leaving the market such that from a welfare point of view it is not clear which regime is preferable. Bloch and Demange (2017) present a setting where consumers are heterogeneous with respect to their privacy cost and a monopolistic platform decides how much data to extract. They show that depending on parameter values the amount of data collection can be excessively high. A similar setting is presented in Lefouili and Toh (2017) where a monopolistic platform monetizes on disclosing personal information to third parties. The authors conclude that one of the inefficiencies arising is excessive information disclosure. The mentioned papers consider the case of monopolistic platforms, while we consider the case of competing platforms, allowing for varying degrees of competition intensity on both market sides.

The role of privacy in a competitive environment is considered in Casadesus-Masanell and Hervas-Drane (2015) where firms not only compete in a price dimension but also in a quality dimension which the authors motivate as privacy. They show that compared to a monopolistic firm, competition leads to a higher degree of privacy while increasing competition intensity does not necessarily imply that privacy improves even further. A key assumption in their model is that prices for disclosing consumer information are exogenous, while in our model platforms have market power vis-à-vis advertisers and hence face a tradeoff. They also show that low privacy firms tend to subsidize consumers, while high privacy firms charge positive consumer prices. Similarly, Kummer and Schulte (2016) show empirically that there is a trade-off between money and privacy for users. They analyze mobile application data and find that apps are cheaper when more personal data can be collected. These results reoccur in our two-sided pricing extension as we show that user prices can be positive or negative as well, while the degree of privacy provision is excessively high or low once firms can no longer compensate users for their data provision. To our knowledge there are very few empirical studies examining the interaction between market power and privacy. In fact, the only study we are aware of is Bonneau and Preibusch (2010) who relate the extent of data collection policies of various online services to the competitiveness of the market they are operating in. They show that the more market power a firm has, the more personal information is asked to be provided which is in line with our model.

3 Model

We analyze a setting where two symmetric platforms, $i, j \in \{1, 2\}$ with $j \neq i$, compete for advertisers and users. Advertisers and users are distributed uniformly on different Hotelling lines of unit length and are assumed to both single-home. This assumption allows us to focus on the role of competition intensity more clearly.[5] Platforms are located at the ends of the respective Hotelling lines such that platform i is located at location $l_i = 0$ and platform j at $l_j = 1$. Note that on the advertiser and the user side we have distinct Hotelling lines and therefore distinct parameters of transportation costs, which we will later interpret as different degrees of competition intensity. The idea is that the degree of competition faced by platforms does not have to be the same for all market sides. For example, online platforms from different segments, such as search engines, social networks, video streaming platforms or mail providers, may all compete for the same advertisers, however competition for users may occur separately and independently of the other segments.

3.1 Users

A user located at x on the Hotelling line obtains utility $u_i(x)$ from joining platform i,

$$u_i(x) = \underline{u} - \kappa(d_i) - \nu(d_i)A_i - t_u|l_i - x|. \tag{1}$$

The first term of the utility function is a fixed utility component $\underline{u}$ from using platform services, which is the same for both platforms. Second, $\kappa(d_i) \geq 0$ denotes the privacy (opportunity) costs of providing user data d_i to the platform, whereby we assume that costs are strictly convex and twice differentiable, and specifically that $\kappa'(0) = 0$, while $\kappa'(d) > 0$ for all $d > 0$ and $\kappa''(d) > 0$ for all d. Third, users incur nuisance cost $\nu(d) \geq 0$ per advertisements A_i on the platform. We assume that users (weakly) prefer personalized to non-personalized ads, i.e. $\nu(d)$ is a convex and twice differentiable function s.t. $\nu'(d) \leq 0$ and $\nu''(d) \geq 0$. This setup reflects the idea that the more relevant an ad, the higher the chance of value creation through a possible follow-up purchase.[6] Finally, users face transportation costs due to horizontal platform differentiation, whereby we

[5] In Section 7 we discuss multi-homing.

[6] Note that our set-up allows for positive utility of seeing advertisement as well, as long as this positive utility is again concave in the amount of provided data. However, for sake of clarity we stay with the notion of negative utility of nuisance in the subsequent text and consider the case of positive cross-group externalities as an extension in Section 7.

assume uniform user distribution on the Hotelling line, i.e. $x \overset{u}{\sim} [0, 1]$, while $t_u > 0$ is the associated transportation cost parameter.

Consumers in our baseline model are not charged a monetary price explicitly, which makes our model comparable to e.g. Reisinger (2012). We follow the same line of reasoning as e.g. in Peitz and Reisinger (2016) and Waehrer (2015) that there are some exogenous constraints preventing platforms from charging non-zero consumer prices. This restriction is, however, relaxed in Section 7.1. In order to join a platform users have to provide some personal data d_i in our model. This is different to the setup in Reisinger (2012) or Bourreau et al. (2017) as in our model platforms can set the level of data which has to be provided by the users, whereas in their models consumers voluntarily provide a certain amount of time. The idea behind our setup is that consumers accept terms and conditions when using a platform which requires them to accept a certain level of data provision or alternatively cases where users have to register for an account by providing personal information before they can use the platform service. This specification on the consumer side allows us to focus on user data d_i as a primary strategic aspect of competition.

3.2 Advertisers

An advertiser located at a on the Hotelling line obtains an expected profit of $\pi_i(a)$ from posting a single ad on platform i,

$$\pi_i(a) = \tau(d_i)(1 - p_i)X_i - t_a|l_i - a|. \tag{2}$$

The interaction with X_i users on platform i generates a normalized expected revenue of 1, if users decide to 'click on the ad', which happens with probability $\tau(d_i)$. The strictly concave and twice differentiable function $\tau(d) \geq 0$ can be interpreted as the targeting ability of platforms: the more data d can be collected from users, the more effective the targeting and hence the higher the probability that a user clicks on this ad, i.e. we have that $\tau'(d) > 0$ and $\tau''(d) < 0$. At the same time we assume that advertisers only pay the platform a price p_i if the ad has been clicked (cost-per-click) such that the expected revenue per user is given by $\tau(d_i)(1 - p_i)$, which is consistent with real-world pricing practices. The second term reflects advertisers transportation costs when joining platform i. Again we assume uniform advertiser distribution on the Hotelling line, i.e. $a \overset{u}{\sim} [0, 1]$, and $t_a > 0$ as the transportation cost parameter on the advertiser side.

3.3 Platforms

The business model of platforms in our model is purely ad-based. While they offer (exogenous) platform services ($\underline{u}$) to users, revenue is only generated through presenting ads to users.[7] Platform profits are then given by

$$\Pi_i\left(d_i, p_i\right) = A_i X_i \tau(d_i) p_i \tag{3}$$

i.e. A_i advertisers at platform i pay p_i whenever the platform's users X_i click on an ad with probability $\tau(d_i)$.[8] The crucial novelty in our model is that we assume that besides charging prices to advertisers, platforms extract data d_i from their users. While d_i shares some price characteristics from the point of view of users, data is an essential input factor for the click-probability the advertisers are facing. At the same time we assume that not only the click probability increases through better targeting possibilities but also the nuisance decreases.

3.4 Assumptions

We make the following assumptions to ensure full advertiser and user market coverage, allowing us to study environments of full platform competition.[9]

Assumption 1. *Competition for advertisers is sufficiently strong, i.e.* $t_a \leq \bar{t}_a$.
This implies that competition for users is sufficiently weak and that there are gains of trade for all advertisers, even without data collection, i.e.

$$(a) \quad t_u > \nu(0),$$
$$(b) \quad t_a < \tau(0).$$

The upper bound on t_a is given by $\bar{t}_a := \frac{t_u \tau(0) - \nu(0)\tau(0)}{3t_u + \nu(0)}$. This assumption on the upper bound of t_a allows us to isolate effects in a competitive environment. Intuitively, this constitutes a sufficient condition, such that for any level of (symmetric) data provision $d \geq 0$, it is assured that all advertisers obtain non-negative profits. Consequently, competition for advertisers is sufficiently strong.

[7] In Section 7 we discuss two-sided pricing.

[8] Note that platforms and advertisers share the profit created by each targeted user on the platform. However, this does not mean that their incentives are perfectly aligned, since platforms additionally care about the number of advertisers joining.

[9] In Section 7 we discuss relaxing the full-market coverage assumptions.

The condition on the consumer nuisance function, i.e. the necessary condition (a) of Assumption 1, can be motivated as follows: no platform will obtain the entire user market, even if all ads were placed on the rival platform. Technically, this is established by $t_u > \nu(0)$.[10] The condition on the targeting technology, i.e. the necessary condition (b) of Assumption 1, states that even without collecting any data advertisers can still profitably join a platform. In particular we assume that there are gains of trade for all advertisers. Intuitively, this assumption states that there is a positive probability for users to click an ad even if the ad is not targeted at all. And this probability, $\tau(0)$, exceeds the transportation cost incurred by any advertiser t_a, so that we need not exclude any advertisers, even if too little data is collected.

Assumption 2. *The fixed utility component $\underline{u}$ is large enough to ensure full participation on the user side.*

Intuitively, the platform service provides sufficient utility such that users are not deterred through the provision of personal data and seeing ads.

The timing of the game is as follows. In the first stage platforms simultaneously set prices p_i and the required level of data d_i to join their platform. In the second stage advertisers and users observe the platforms' choices and simultaneously decide which platform to join, hence determining A_i and X_i.[11] The equilibrium concept is sub-game perfection and we solve the game by backward induction.

4 Equilibrium Analysis

In this section we will first present the results for the second-stage sub-game of user and advertiser allocation. Then we will show the efficient and the user-optimal outcome as well as the market outcome in the Sub-game Perfect Nash Equilibrium.

4.1 Second Stage Market Shares

In the second stage the market shares on the consumer and advertiser side are given by the standard Hotelling procedure. Utilizing the unit length of the Hotelling line,

[10] Note that $t_u > \nu(0) \Rightarrow t_u > \nu(d)\,\forall d$ because $\nu'(d) \leq 0$. Given any (symmetric) amount of data $d \geq 0$ collected by both platforms, even if all advertisers used platform j such that $A_i = 0$ and $A_i = 1$, at least the user most loyal to platform j, i.e. located directly at l_j, would rather stay at this platform j, even though it is full of ads. In other words, competition for users is sufficiently weak.

[11] We could also consider an alternative timing where advertisers choose first and users last. The outcome is equivalent in our model.

and given full user market coverage due to Assumption 2, the number of users joining a platform is then determined by the indifferent consumer $\hat{x} : u_i(\hat{x}) = u_j(\hat{x})$ such that

$$X_i = \hat{x} = \frac{1}{2} + \frac{1}{2t_u}\left[\kappa(d_j) - \kappa(d_i) + \nu(d_j)A_j - \nu(d_i)A_i\right], \qquad X_j = 1 - \hat{x}. \tag{4}$$

Similarly, market shares on the advertiser side are given by the indifferent advertiser $\hat{a} : \pi_i(\hat{a}) = \pi_j(\hat{a})$. Note that Assumption 1 assures market coverage gross of advertising prices. For now we therefore assume that prices permit full market coverage and check later that in equilibrium this is indeed the case. Market shares are then given by

$$A_i = \hat{a} = \frac{1}{2} + \frac{1}{2t_a}\left[\tau(d_i)(1 - p_i)X_i - \tau(d_j)(1 - p_j)X_j\right], \qquad A_j = 1 - \hat{a}. \tag{5}$$

Solving the system of equations given in (4) - (5) yields unique market shares X_i, X_j, A_i and A_j as functions of data requirements d_i, d_j and prices p_i, p_j. Explicit solutions are provided in the Appendix.

4.2 Efficiency Benchmark

For the derivation of the welfare-efficient benchmark, we define welfare as the sum of all indirect utilities and profits, anticipating second stage market shares as in 4.1, i.e.

$$W\left(d_i, d_j, p_i, p_j\right) = \int_0^{X_i} u_i(x)\mathrm{d}x + \int_{X_i}^1 u_j(x)\mathrm{d}x + \int_0^{A_i} \pi_i(a)\mathrm{d}a + \int_{A_i}^1 \pi_j(a)\mathrm{d}a + \Pi_i + \Pi_j. \tag{6}$$

Proposition 1. *Welfare is maximized by the unique symmetric solution* $(d^o, p^o) = (d_i^o, p_i^o)$ *for* $i \in \{1, 2\}$*, where* d^o *is characterized by*

$$\kappa'(d^o) = \frac{\tau'(d^o)}{2} - \frac{\nu'(d^o)}{2} \tag{7}$$

resulting in equal advertiser and user market shares, i.e. $A_i^o = 1/2$ *and* $X_i^o = 1/2$*. The price* p^o *can be freely chosen to split the rent between advertisers and platforms.*

The welfare-optimal level of data d^o is chosen in a way such that users' marginal cost of data provision $\kappa'(d^o)$ equals the sum of marginal benefits across both market sides, i.e. the marginal benefit of enhanced targeting $\tau'(d^o)/2$ and the marginal benefit of reduced nuisance $-\nu'(d^o)/2$, while the factor $1/2$ is due to the symmetric market shares.[12] Furthermore, the optimal level of data provision is independent of transportation cost parameters t_a and t_u. Since prices are just transfers from advertisers to platforms they do not affect welfare.[13]

4.3 User-optimal Outcome

Let us now turn to the user-optimal level of data provision. If users are free to decide on the amount of data provided, the user-optimal level d^u is derived from consumer surplus, which is identical to the first two terms in equation (6), anticipating second stage market shares as in 4.1.[14]

Proposition 2. *User utility is maximized by the unique symmetric solution* $(d^u, p^u) = (d_i^u, p_i^u)$ *for* $i \in \{1, 2\}$, *where* d^u *is characterized by*

$$\kappa'(d^u) = -\frac{1}{2}\, \nu'(d^u), \tag{8}$$

while the price p^u *can be freely chosen to split the rent between advertisers and platforms, resulting in equal advertiser and user market shares, i.e.* $A_i^u = 1/2$ *and* $X_i^u = 1/2$.

Intuitively, the user-optimal data level balances privacy costs and reduced nuisance benefits for users, at the margin. Note that for constant nuisance costs we get the corner-solution where users would not provide any private data, i.e. $d^u = 0$. For general decreasing nuisance costs, users would be willing to provide a positive level of data $d^u > 0$.

4.4 Market Outcome

For the market outcome, in the first stage platforms maximize their profits, anticipating second stage market shares as in Section 4.1.

$$\max_{p_i, d_i} \Pi_i\,(d_i, p_i) = A_i\, \tau(d_i)\, p_i\, X_i \;\forall i \in \{1, 2\} \tag{9}$$

[12] For very low transportation cost parameters and sufficiently high net benefits $\tau(\cdot) - \nu(\cdot)$ on the platform it might be efficient from a welfare perspective to shut one platform down and let the entire market be served by the other platform due to high network effects. In this case the very fact of having a competing platform is an inefficiency. While this corner solution exhibits an interesting property of platform markets, it is not the focus of this paper and we therefore stick to the case where we have an interior, i.e. duopoly solution as the efficient benchmark.

[13] The same data level d^o would result if we only choose d_i to maximize welfare, while anticipating firms setting ad prices p_i subsequently. These prices would be identical to the prices in the market outcome, given by equation (13). The same argument applies for the user optimal level d^u.

[14] See footnote 13.

We obtain first-order conditions for prices and data levels, i.e.

$$\frac{\tau'(d_i)}{\tau(d_i)} = - \frac{\frac{\partial A_i}{\partial d_i} X_i + \frac{\partial X_i}{\partial d_i} A_i}{A_i \, X_i}, \tag{10}$$

$$p_i = \frac{A_i \, X_i}{\frac{\partial A_i}{\partial p_i} X_i + \frac{\partial X_i}{\partial p_i} A_i}. \tag{11}$$

Intuitively, equation (10) says that targeting benefits of data collection must equal the effects on user and advertiser shares, at the margin. Similarly, also prices (11) must reflect their impact on user and advertiser shares. Regarding the curvature of the maximization problem we note that the solution to the first-order conditions represents a maximum as long as the targeting technology $\tau(\cdot)$ is sufficiently concave, the nuisance cost $\nu(\cdot)$ is sufficiently convex, or both. The details of this condition are given in Appendix A.

Solving the set of first-order conditions we obtain the following symmetric solutions for prices and data levels.

Proposition 3. *There exists a (symmetric) Sub-game Perfect Nash Equilibrium with* $(d_i^*, p_i^*) = (d^*, p^*)$ *for* $i \in \{1, 2\}$*, such that the level of data collected from a users is implicitly given by*

$$\kappa'(d^*) = \left(\frac{\nu(d^*) + t_u}{\tau(d^*) - t_a} \right) \frac{\tau'(d^*)}{2} - \frac{\nu'(d^*)}{2} \tag{12}$$

and prices per advertisement are

$$p^* = 2 \frac{t_a t_u + \nu(d^*) \tau(d^*)}{\tau(d^*) \left[t_u + \nu(d^*) \right]}, \tag{13}$$

resulting in equal advertiser and user market shares, i.e. $A_i^* = 1/2$ *and* $X_i^* = 1/2$.

Comparing the market level of data provision d^* in (12) to the efficient level d^o in (7) we see that the marginal targeting benefit $\frac{\tau'(d^*)}{2}$ is additionally weighted by $\frac{\nu(d^*)+t_u}{\tau(d^*)-t_a}$. This distortion is analyzed in detail in Section 6. Note that the equilibrium price p^* does not exceed one and that profits are positive for all advertisers due to Assumption 1.[15]

Before we continue we state a corollary concerning the equilibrium effect of data provision on user utility.

Corollary 1. *In equilibrium,* $\kappa'(d^*) > -\nu'(d^*)/2$.

[15] In Appendix A we provide the details for this result.

Proof. See Appendix A.4. □

Intuitively, Corollary 1 implies that in equilibrium users' data provision is such that the (negative) privacy costs effect on user utility is larger than the (positive) effect of reduced nuisance. Consequently, in the market outcome too much personal data is provided compared to the user-optimal level.[16]

5 Comparative Statics

In this section we want to provide economic intuition for the equilibrium results of our model. For this we will provide comparative statics, given changes in advertiser-side competition intensity t_a and user-side competition intensity t_u as well as nuisance $\nu(d)$ and targeting $\tau(d)$ on equilibrium values of personal data provision d^*, ad-per-click price p^*, as well as platform profits Π_i^*, advertiser profits π_i^* and user utility u_i^*.

As most of the comparative statics effects are in line with standard intuition from two-sided platforms, we delegate these analyses to the Online Appendix B and refer to the table in Figure 1 for an overview of all derived comparative statics results. In this section we focus on the important and seemingly counter-intuitive effects of competition intensities of both market side.

FIGURE 1: Overview of Comparative Statics

z	$\mathrm{d}d^*/\mathrm{d}z$	$\mathrm{d}p^*/\mathrm{d}z$	$\mathrm{d}\Pi_i^*/\mathrm{d}z$	$\mathrm{d}\pi_i^*/\mathrm{d}z$	$\mathrm{d}u_i^*/\mathrm{d}z$
t_a	+	+	+	−	−
t_u	+	−	−	+	−
$\nu(d)$	+	+	+	−	−
$\tau(d)$	−	?	+	?	+

Note that we distinguish between the platform competition intensity on the user side and on the advertiser side. As platforms are horizontally differentiated vis-à-vis both market sides, competition intensity on each side can be measured through the corresponding transportation cost parameter: higher transportation costs mean higher platform differentiation and thus higher switching costs on this market side, which can be interpreted as more platform market power and hence lower competition intensity.

[16] In Section 6 we provide a detailed comparison of the market outcome and the user-optimal outcome.

5.1 Advertiser-side Competition

First, we consider the effects of advertiser-side competition on data collection. For this consider the platform's first-order condition in equation (10) and note that the data level choice depends on the effects of d_i on advertiser and user market shares A_i and X_i. Regarding market share reactions we obtain $\partial X_i/\partial d_i < 0$ and $\partial A_i/\partial d_i < 0$ at equilibrium values.[17] Intuitively, additional data provision d_i would shy away users X_i because marginal privacy costs are higher than marginal benefits of reduced nuisance (compare Corollary 1). Although more data provision increases targeting, overall, advertisers would still be repelled by additional data provision because of the detrimental effect on user market share at that platform.

In equilibrium, if competition for advertisers softens, i.e. transportation costs t_a increase, advertisers become 'more sticky', i.e. less sensitive to changes in data provision (and hence user demand) such that $\partial^2 A_i/(\partial d_i \partial t_a) > 0$. Contrary, users become more sensitive to data provision such that $\partial^2 X_i/(\partial d_i \partial t_a) < 0$. Overall, the former effect dominates the latter effect in magnitude. Consequently, and recalling $X_i^* = A_i^* = 1/2$, the right-hand-side of equation (10) decreases in t_a such that the equilibrium level of data provision must increase as the left-hand-side is falling in d_i, i.e.

$$\frac{\mathrm{d}d^*}{\mathrm{d}t_a} > 0. \tag{14}$$

This effect might seem counter-intuitive initially. However note that in equilibrium platforms balance the following trade-off for the data level. On the one hand, more data collection yields higher targeting rates, higher advertiser demand and in sum higher profits. On the other hand, collecting more data decreases user demand, which in turn repels advertisers and thus decreases platform profits. If competition for advertisers softens, the latter effect is dampened more than the former effect is strengthened. This yields a new balance of the trade-off, where more user data is collected.

While advertiser prices p^* rise in t_a (compare Online Appendix B), the effect on user data collection d^* does not follow 'standard' two-sided platform logic as here less competition for advertisers, i.e. less sensitive advertiser demand, *increases* users' data 'payment'. Therefore, users actually benefit from increased competition on the advertiser side, such that also $\mathrm{d}u_i^*/\mathrm{d}t_a < 0$, as discussed in the Online Appendix B. Also, since $\mathrm{d}d^*/\mathrm{d}t_a > 0$ and $\mathrm{d}p^*/\mathrm{d}t_a > 0$ we naturally have $\mathrm{d}\Pi_i^*/\mathrm{d}t_a > 0$.

[17] Note that derivations can be found in Appendix A.5.

5.2 User-side Competition

Second, we evaluate the effects of user-side competition intensity on data collection. Similar to the analysis above, we know that $\partial X_i/\partial d_i < 0$ and $\partial A_i/\partial d_i < 0$ in equilibrium. If competition for users softens, i.e. transportation costs t_u increase, on the one side users become less sensitive to changes in data provision such that $\partial^2 X_i/(\partial d_i \partial t_u) > 0$. Therefore, advertisers also become less sensitive to data provision such that $\partial^2 A_i/(\partial d_i \partial t_u) > 0$ because they care about the share of users on that platforms. Therefore the right-hand-side of equation (10) decreases in t_u such that the equilibrium level of data provision must increase, i.e.

$$\frac{\mathrm{d}d^*}{\mathrm{d}t_u} > 0. \tag{15}$$

Two effects are intuitively relevant here. On the one hand, platforms care about the share of users on their platform because it increases their profits directly, but also indirectly through more attracted advertisers. On the other hand, platforms want to increase the level of user data collected as it enhances targeting, attracts advertisers and hence increases profits. In equilibrium, stronger competition for users impacts the former effect of attracting users more than the latter of increasing targeting, therefore, platforms will collect less user data. Following the same intuition, platforms would be willing to lose some advertisers in order to not repel valuable users. Hence, also equilibrium advertiser prices increase in t_u (compare Online Appendix B.1). Contrary to the effects of advertiser-side competition, these results reflect the 'standard' two-sided platform logic: stronger competition for users reduces the 'price' on the user side, while it increases the price on the advertiser side.

Furthermore, we discuss the effect of user-side competition intensity on platform profits. One could expect that platforms' profit increases if competition for users becomes less intense, however the opposite is true. Note that their profit function in equilibrium, $\Pi_i^* = p^*\tau(d^*)\,A_i^*\,X_i^* = (1/4)\,p^*\tau(d^*)$. A change in user-side competition intensity t_u gives

$$\frac{\mathrm{d}\Pi_i^*}{\mathrm{d}t_u} = \frac{1}{4}\left[\frac{\mathrm{d}p^*}{\mathrm{d}t_u}\tau(d^*) + \tau'(d^*)\frac{\mathrm{d}d^*}{\mathrm{d}t_u}p^*\right]. \tag{16}$$

On the one hand, advertiser prices decrease if competition for users becomes less intense (t_u increases), which reduces platform profits. Hence the first term on the right-hand side of (16) is negative. On the other hand, the second term is positive, because when competition for users becomes less intense (t_u increases), more data can be collected from users, which leads to more effective ad targeting and therefore increased platform profits. As can be seen from the derivation in Appendix A, overall, the negative first-term effect is stronger in equilibrium, such that platforms suffer from weaker competition for users, i.e. $\mathrm{d}\Pi_i^*/\mathrm{d}t_u < 0$.

6 Policy Implications

In this section we draw comparisons between the different outcomes outlined in Section 4 and present policy implications.

6.1 Comparison of Outcomes

First, we want to compare the outcome of the efficiency benchmark with the market equilibrium outcome. If we compare the right-hand-side of the competitive level d^* in (12) and the efficient level d^o in (7) we can see that the difference will crucially depend on the distortion induced by

$$\delta(d^*) := \frac{\nu(d^*) + t_u}{\tau(d^*) - t_a}, \tag{17}$$

which gives more or less weight to the marginal benefit on the advertiser market side $\tau'(d^*)/2$. Note that by Assumption 1 the denominator of $\delta(d^*)$ is positive, so that we have $\delta(d^*) > 0$. As the efficient level d^o does not depend on parameter values, we can see that there can be underprovision ($d_u^* < d^o$) as well as overprovision ($d_o^* > d^o$) of personal data in the competitive equilibrium. Depending on the structure of the market too much or too little weight is put on the advertiser side of the market. In particular we can infer from equations (12) and (7) that the competitive outcome leads to underprovision of personal data if $\delta(d^*) < 1$ and to overprovision if $\delta(d^*) > 1$. Note for $\delta(d^*) = 1$ expression (12) simplifies to (7), the efficient level of data provision. Using our definition of $\delta(d^*)$ we can then see that $d^* < d^o$ if

$$\delta(d^*) < 1 \iff \tau(d^*) - \nu(d^*) > t_a + t_u \tag{18}$$

and $d^* > d^o$ if

$$\delta(d^*) > 1 \iff \tau(d^*) - \nu(d^*) < t_a + t_u. \tag{19}$$

These results are summarized in the following proposition.

Proposition 4. *The competitive outcome leads to overprovision of personal data if competition on both market sides is weak and/or if net cross-group externalities are small. If competition on both market sides is strong and/or net cross-group externalities are large, the competitive outcome exhibits underprovision of personal data.*

Proof. See Appendix A.4. □

We want to interpret this finding by first holding the functions $\kappa(d)$, $\nu(d)$ and $\tau(d)$ fixed and asking the question which competitive environment leads to which scenario. From our comparative statics results we know that the amount of data is a monotone function of the transportation cost parameters, i.e. $\frac{\mathrm{d}d^*}{\mathrm{d}t_u} > 0$ and $\frac{\mathrm{d}d^*}{\mathrm{d}t_a} > 0$. Proposition 4 then gives us a threshold for how the resulting level of data collection compares to the efficient benchmark: if competition is too strong, i.e. $t_a + t_u$ is small, platforms tend to collect and process an inefficiently small amount of data as users and advertisers shy away too easily. If in turn competition on both sides is weak, i.e. $t_a + t_u$ is high, the market sides become more sticky and platforms are able to extract high amounts of personal data.

We can also hold the competitive environment t_a, t_u on both sides fixed and analyze the effects of relatively strong or weak opposing cross-group externalities. On the one hand, an additional user imposes a positive externality on advertisers (and platforms), which is equal to the targeting effect $\tau(d^*)$. On the other hand, an additional advertiser imposes a negative externality on users, which is equal to the nuisance costs $-\nu(d^*)$. The net effect can therefore be interpreted as available gains from trade in this economy. If the net effect is relatively large, there are significant gains of trade which could be seized by increasing the amount of data collected. If the net effect is small, the gains from trade could be increased by lowering the amount of collected data.

Comparing the user-optimal level d^u to the welfare-optimal level d^o we immediately see that users would provide an inefficiently low level of data. This result is summarized in the following proposition.

Proposition 5. *The user-optimal level of data provision is inefficiently low.*

The reason for this result is straightforward. As users do not internalize the effect the data has on the advertiser market, they will provide data up to the point where the marginal decrease in nuisance equals marginal cost of data provision. Since from a welfare perspective the value creation aspect on the advertiser market is omitted, the resulting level of data provision is inefficiently low. Furthermore, since $\delta(d^*) > 0$ we also have $d^* > d^u$ for all exogenous parameters and functional forms, as shown in Corollary 1. Unlike users, platforms act as intermediaries and are able to internalize parts of the value creation on both sides of the market.

6.2 Policy Conclusions

In this section we briefly discuss what conclusions can be drawn from our previous analyses when it comes to policy implications and regulation.

In our model, an omnipotent regulator could obviously achieve the first-best outcome by forcing $d_i = d_j = d^o$ and increasing competition on both sides of the market such that $t_u \to 0$ and $t_a \to 0$. In this case the efficient amount of data is provided while the total transportation costs approach zero.

In practice, regulation and policy discussions typically focus on data and privacy regulation or on competition policy measures (or merger regulation) to assure competitiveness on the user side, for example in the recent Facebook case at the BKartA or the Facebook/Whatsapp merger case in the US and the EU. In this section we want to present answers our model provides for privacy and competition policy, taking into account both market sides and at the same time the effect on privacy.

Privacy Regulation

Holding the competitive structure of the market fixed, the regulator could improve upon the market outcome by enforcing the efficient level of private data provision $d_i = d_j = d^o$. However, a direct regulation of the amount of data in our model requires knowledge of the cross-group externalities, i.e. functions $\tau(d)$ and $\nu(d)$, as well as users' privacy concerns $\kappa(d)$.

A regulator could also consider switching to a consumer standard and let consumer freely choose how much data they would like to provide. Our results show that the user-optimal amount of data is always inefficiently low as users do not internalize the benefit on the advertiser side. In particular our results suggest that we can only improve in terms of welfare by switching to a consumer standard when there is extreme overprovision of data in the economy, i.e. platforms have significant market power on both sides of the market. If the market exhibits underprovision, switching to the consumer standard always reduces welfare.

Competition Policy

An approach which is less demanding when it comes to information requirements is the regulation of the competitive environment on both market sides, i.e. t_u and t_a. Our results (Proposition 4) suggest that if competition is very weak on both sides ($t_u + t_a$ high), the amount of data collected is likely to be inefficiently high. Similarly, if competition is too strong ($t_u + t_a$ low), too little data is provided from a welfare point of view. While regulators still have to know whether there is overprovision or underprovision in the market in the first place, our results can still provide some guidance.

Our comparative statics results suggest that increasing competition works in the same direction for both sides of the market. The equilibrium amount of data provision is a monotone function of the transportation cost parameters t_a and t_u and by altering either one of the parameters it is possible to push the competitive equilibrium amount of data d^* towards the welfare optimum d^o. Typical examples include reducing switching costs on the user side (see e.g. GDPR/data portability in the EU) or policing vertical integration on the advertiser side (see e.g. debate around Google/DoubleClick acquisition). Further, our results suggest that more competition between platforms is not necessarily welfare enhancing as it further limits the ability to create economic value through the collection of personal data in the case of underprovision.

Also, our results suggest that policy measures, although they work in the same direction, are not equally effective across market sides, i.e. $\frac{\mathrm{d}d^*}{\mathrm{d}t_a} \neq \frac{\mathrm{d}d^*}{\mathrm{d}t_u}$. This might be particularly important in a scenario where the market exhibits underprovision and a regulator would have to reduce competition as this implies increasing transportation costs in the economy. Increasing transportation costs would then lead to more data collection in the subsequent market outcome. Whether we can increase total welfare by increasing transportation costs, however, depends crucially on whether the benefit of higher and thus more efficient data provision (non linear) exceeds the increased costs of transportation (linear).[18] This trade-off could call for a second-best regulation, where competition intensity is regulated in such a way that the amount of data provided in the subsequent market outcome balances the above mentioned benefits and costs at the margin.

From these results on competition policy we want to draw two main conclusions. First, regulating competition on either or both market sides can address the privacy / data collection distortion in the market outcome. Second, whenever regulators consider

[18]Note that also in a situation of overprovision, the market structure might be such that it is socially beneficial to decrease transportation costs, i.e. increase competition, even beyond the level where it induces efficient data provision (as established in equation 7), such that the benefits of decreased transportation costs outweigh the costs from data underprovision.

competition policy or merger regulation in these data-driven industries, they should take into account the impact on data collection in the market.

7 Discussion

In this section we sketch and briefly discuss extensions and variations of the baseline model presented in Section 3.

7.1 User Prices

In this section we consider an alternative setup where platforms can charge prices on the user side of the market, too. All other model specifications remain as before, i.e. specifically users now have to pay a monetary price additional to their personal data 'payment'. In a sense, this setup could be considered as an unrestricted model, where platforms are not restricted to zero user prices. Let p_i^u denote the price a user has to pay to join platform i. User utility is then given by

$$u_i(x) = \underline{v}_i + \underline{d} - \kappa(d_i) - \nu(d_i)A_i - p_i^u - t_c|l_i - x|, \tag{20}$$

while advertisers still face the same decision as in Section 3. Market shares are obtained as before by pinning down indifferent users and advertisers and solving the resulting system of equations. The resulting profit maximization problem of platform i is then given by

$$\max_{p_i, d_i, p_i^c} = A_i\tau(d_i)p_iX_i + p_i^uX_i \ \forall i \in \{1, 2\}. \tag{21}$$

Following the same procedure as in our baseline model we obtain symmetric equilibrium values $p_i = p_j = \tilde{p}$, $p_i^u = p_j^u = \tilde{p}^u$ and $d_i = d_j = \tilde{d}$ where advertiser prices are given by $\tilde{p} = 2[t_a + \nu(\tilde{d})]/\tau(\tilde{d})$, user prices by

$$\tilde{p}^u = t_a + t_c + \nu(\tilde{d}) - \tau(\tilde{d}), \tag{22}$$

while the equilibrium amount of data is given by

$$\kappa'(\tilde{d}) = \frac{1}{2}\left[\tau'(\tilde{d}) - \nu'(\tilde{d})\right]. \tag{23}$$

We immediately see from equations (7) and (23) that $\tilde{d} = d^o$.

Proposition 6. *If platforms can charge prices on both market sides, the efficient level of data is collected.*

Since platforms can now extract rents from both sides of the market, they maximize the aggregate value, whereas in our baseline model platforms only profited on the advertiser side of the market and hence set a data requirement level which is distorted. Taking a closer look at equilibrium user prices in (22) we immediately see that negative, positive or zero user prices are possible, depending on parameter values and functional forms.

Proposition 7. *If user prices in the two-sided pricing model are positive, the one-sided pricing constraint would result in data overprovision. Contrary, if user prices are negative, this constraint would yield underprovision.*

Proof. See Appendix A.4. □

The intuition for this result is that now platforms can extract the efficient amount of data by adequately compensating users. If net benefits of data collection are large or competition is rather strong, platforms can extract large amounts of data from users and then compensate them by charging negative user prices, whereas in the one-sided pricing model platforms do not have the instrument for compensation and therefore are forced to collect less data than the efficient level. Vice versa, if net benefits are small or competition rather weak, platforms are not forced to monetize through ads by extracting an inefficiently high amount of data, but can obtain positive revenue from the user side instead and leave the amount of data at the efficient level.

We would like to mention at this point that this result may depend on the fact that even with positive user prices we assume the user market to remain fully covered. However, remember that under a market solution with overprovision users gain in terms of utility by decreasing d from d^* to d^o. If this difference in utility is enough to cover the associated positive user price, the user market remains covered. If the consumer price exceeds the utility gain, the two-sided pricing may lead to users leaving the market and efficiency may not be feasible any longer. We provide a more detailed discussion of the full market coverage assumption in the subsequent section. A similar argument can also be made if we consider heterogeneous users as then our uniform pricing setup may not be sufficient to ensure efficiency but platforms would need to engage in price discrimination.

Nevertheless, we would like to draw two further conclusions from these results. Firstly, observing a user price $\tilde{p}^u = 0$ empirically is consistent with the equilibrium result above as well as with our baseline model. By observing zero prices we can not infer whether a price of zero is an optimal choice, making the model above the 'correct' model, or whether there are constraints which prevent platforms from setting user prices at all, making our baseline model more suitable. Secondly, since user prices depend on parameters of competition intensity and externalities, observing zero prices across different markets, jurisdictions and industry sectors makes it unlikely that $\tilde{p}^u = 0$ is a profit maximizing choice in all cases. This strongly supports the argument made by Waehrer (2015) that user prices are not a (practical) variable of interest in real-world platform maximization problems.

7.2 Collusion

Full Collusion

Let us consider a collusive game where platforms agree on prices $p_i = p_j = p$ and data requirements $d_i = d_j = d$ such that joint profits are maximized. Since advertisers face transportation costs, the profit maximizing collusive price is such that the participation constraint of the indifferent advertiser is binding $\pi_i\left(\frac{1}{2}\right) = 0$ which yields $p = 1 - \frac{t_a}{\tau(d)}$. Plugging the collusive price p into the platforms' profit functions (3) we obtain $\Pi_i = \frac{1}{4}(\tau(d) - t_a)$ and immediately see that profits are increasing in d up to the point where the participation constraint of the indifferent user binds $d : u_i(\frac{1}{2}) = 0$. Since we assumed $\underline{u}$ to be high enough to have interior solutions in the previous sections, we can infer that the collusive amount of data will be excessively high.

Partial Collusion

In this section we consider an alternative collusive environment where platforms coordinate on setting a symmetric level of data d but still compete in prices on the advertiser market. The idea is that platforms might influence privacy regulation in a collusive effort without coordinating their pricing decisions. We therefore introduce a collusive stage where platforms agree on a symmetric level d prior to the price setting decision. It is easy to verify that symmetric prices are then given by $p_i = p_j = p(d) \equiv 2\frac{t_a t_u + \nu(d)\tau(d)}{\tau(d)[t_u + \nu(d)]}$, similar to the market outcome outlined in Section 4. The key difference, however, is the collusive choice of d. As prices (and d) are symmetric, market shares can be anticipated to be given by $A_i = A_j = X_i = X_j = 1/2$ such that industry wide platform profits are given by $\Pi(d) := \Pi_i(d) + \Pi_j(d) = \frac{\tau(d)p(d)}{2}$.

If we have $\Pi'(d) > 0$ for all d, the collusive level will be the same as in full collusion case, such that the participation constraint of the users will be binding, and if $\Pi'(d) = 0$ has a solution, a possible interior solution exists. The comparison to the market outcome (or to the efficient outcome) is in this case, however, ambiguous and depends on functional forms and parameter values.

Interestingly, industry profits are not necessarily increasing in d. In fact if $\Pi'(d) = p(d)\tau'(d) + p'(d)\tau(d) < 0$ for all d then the collusive level of data will be zero. The reason for this seemingly counter-intuitive result is that increasing d can effectively propagate competition on the advertiser market. In particular if we go back to the definition of advertiser market shares in (5) we can see that increasing a symmetric level d has the same effect on the advertiser market as a decrease in transportation costs in the sense that it makes advertisers more reactive towards changes in prices. The intuition is straightforward: if the click-probability is very high, small differences in prices become magnified. The trade-off faced by the platforms is then the following. An increase in click probability (through increasing d) results in tighter competition on the ad market (depressing p). The optimal d can therefore vary widely depending on which effect dominates.

To briefly summarize this section we can conclude that full collusion amongst platforms should be avoided whenever possible. When it comes to partial collusion, however, a more nuanced analysis is necessary as competition on the ad market might be sufficiently strong to prevent inefficient regulatory capture.

7.3 Market Coverage and Multi-homing

In this section we want to briefly discuss the effects of relaxing the assumptions guaranteeing full market coverage and single-homing. We consider market-coverage and multi-homing together because without these assumptions in both cases the market share of a platform is determined by the user/advertiser who is indifferent between joining a platform and the outside option, whereas in the baseline model it was determined by the user/advertiser who is indifferent between joining both platforms. Note that this changes the interpretation of transportation costs in the model substantially. While in the baseline model transportation costs measured a restriction to switching to the other platform and hence a degree of platform competition, now they rather exhibit a restraint on a platform's demand, independent of the other platform. Essentially, lower transportation costs can now be interpreted as *more* elastic demand, whereas in the baseline model they reflected *less* elastic (sticky) demand. While our assumptions for the baseline model were chosen to study full competition between platforms, relaxing the assumptions on one market side significantly changes the setting in the sense that platforms now only compete indirectly through the other market side. Nevertheless, we want to provide some intuition for the robustness of our results. For a more detailed analysis consider the Online Appendix B.

Advertiser Side

On the advertiser side, lifting Assumption 1 of a covered market together with the single-homing assumption can result in two cases, depending on parameters. First, if transportation costs t_a are sufficiently small, some advertisers 'in the middle' will use both platform (multi-homing). The comparison of the new equilibrium level of data provision to the new efficient level or the baseline level of data provision is, however, ambiguous. This is because less advertiser demand elasticity on the one hand could allow firms to readjust d, while at the same time the total number of advertisers on a platform could rise. From an efficiency perspective, though, more data should be collected than was efficient in the baseline model. Second, if transportation costs t_a are sufficiently high, some advertisers in the middle might choose not to use any platform (no full market coverage). Then it would also be efficient to exclude some advertisers such that the new efficient level of data provision is below the efficient baseline level. The comparison to the equilibrium outcomes remains however ambiguous, as above.

User Side

On the user side, relaxing the full-market Assumption 2 and the single-homing constraint similarly leads to either some users 'in the middle' joining both platforms (multi-homing) or some user joining neither platform (no full market coverage), depending mainly on transportation costs t_u. In both cases user demand is then merely scaled by the demand elasticity, i.e. the transportation costs t_u, and users' role essentially reduces to being a resource of data needed to create advertising surplus.[19] We find that there would always be over-provision of user data in equilibrium because the efficient benchmark takes into account the trade-off between total value creation and user exclusion, whereas the market outcome only balances targeting benefits and potential user exclusion. However, still less data is collected than in the baseline model and also the efficient level of data decreases. Further, we find that now the transportation cost parameters have no effect on the

[19]Note that on the advertiser side this was not the case because advertisers pay money rather than a value-creating resource.

equilibrium level of data provision. This is because t_u merely scales demand while the relevant trade-off for the choice of d involves the actual utility when joining the platform and is not influenced by the demand scale. Furthermore, equilibrium prices now increase in t_u and decrease in t_a. Because of the reversed role transportation costs now play, this is not contradictory to the baseline model results: the harder it is to keep users, the higher the price for advertisers. Consequently, platform profits still increase and advertiser profits still decrease in user-side elasticity.

7.4 Positive Cross-group Externalities

In the baseline model we considered the case where users incur nuisance cost from seeing ads on the platform, i.e. a negative cross-group externality incurred by users. As explained in the beginning we consider this case because we think it illustrates the main results in a very intuitive way. What we demonstrate in the Online Appendix B is that the model can in fact be generalized to have positive cross-group effects in both directions while the major results remain unchanged.

8 Conclusion

We analyze the role of competition intensity in a two-sided market framework where platforms collect data from users and monetize through ad-sales. Our model predicts that the equilibrium amount of collected data will be distorted compared to the welfare efficient benchmark. Depending on the net effect of cross-group externalities and the competition intensity on both sides of the market, the distortion can lead to underprovision or overprovision of personal data. Since the level of collected data increases the more market power platforms have on either side of the market, side specific regulations are substitutable. We also show that a consumer standard would always lead to underprovision of data as users do not internalize improvements in the targeting capabilities. Lastly, we showed that two-sided pricing induces platforms to choose the efficient level of data by adequately compensating users.

While we think our model provides useful insights we would also like to discuss some limitations. It would be interesting to further explore the role of multi-homing on the advertiser side as it changes the competitive dynamics substantially. Secondly, one could alter the setting on the user side and consider heterogeneous users, while platforms engage in second degree discrimination by offering a menu of data choices. We think those are interesting avenues for future research.

Appendix

A Omitted Oroofs

A.1 Second Stage Market Shares

Note that equations (4) - (5) are consistent, non-redundant and linear in X_i, X_j, A_i, A_j such that the resulting solution in (4.1) is unique. Explicit market shares are then given by:

$$X_i = \frac{t_a(2\kappa(d_j) - 2\kappa(d_i) + \nu(d_j) - \nu(d_i) + 2t_u) + (1-p_j)\tau(d_j)(\nu(d_i) + \nu(d_j))}{4t_a t_u + (\nu(d_i) + \nu(d_j))((1-p_i)\tau(d_i) + (1-p_j)\tau(d_j))}$$

$$X_j = \frac{t_a(2\kappa(d_i) - 2\kappa(d_j) + \nu(d_i) - \nu(d_j) + 2t_u) + (1-p_i)\tau(d_i)(\nu(d_i) + \nu(d_j))}{4t_a t_u + (\nu(d_i) + \nu(d_j))((1-p_i)\tau(d_i) + (1-p_j)\tau(d_j))}$$

$$A_i = \frac{(1-p_i)\tau(d_i)(\kappa(d_j) - \kappa(d_i) + \nu(d_j) + t_u) - (1-p_j)\tau(d_j)(\kappa(d_i) - \kappa(d_j) - \nu(d_j) + t_u) + 2t_a t_u}{4t_a t_u + (\nu(d_i) + \nu(d_j))((1-p_i)\tau(d_i) + (1-p_j)\tau(d_j))}$$

$$A_j = 1 - \frac{(1-p_i)\tau(d_i)(\kappa(d_j) - \kappa(d_i) + \nu(d_j) + t_u) - (1-p_j)\tau(d_j)(\kappa(d_i) - \kappa(d_j) - \nu(d_j) + t_u) + 2t_a t_u}{4t_a t_u + (\nu(d_i) + \nu(d_j))((1-p_i)\tau(d_i) + (1-p_j)\tau(d_j))}$$

A.2 Second-Order Conditions

In the following we derive sufficient conditions such that the equilibrium values p^*, d^* derived from the maximization problem presented in Section 3 characterize a local maximum. Let us consider the Hessian evaluated at equilibrium values. Starting with

$$\left.\frac{\partial^2 \Pi_i}{\partial p_i^2}\right|_{d^*,p^*} = -\frac{t_u^2\ \tau(d^*)^2(\nu(d^*) + t_u)}{4(t_u - \nu(d^*))^2(\nu(d^*)\tau(d^*) + t_a t_u)}$$

we immediately see that $\left.\frac{\partial^2 \Pi_i}{\partial p_i^2}\right|_{d^*,p^*} < 0$, a necessary condition for the Hessian to be negative definite. In the next steps we argue that we can always find functions $\tau(\cdot), \nu(\cdot)$ such that $det(H)|_{d^*,p^*} > 0$. First, it is helpful to look at the numerator and the denominator of the Hessian separately

$$det(H)|_{d^*,p^*} = \frac{H_{num}}{H_{den}}$$

where the numerator H_{num} and the denominator H_{den} are given by

$$\begin{aligned} H_{num} = \tau(d^*)^2 \Big[&-4t_u^2(t_a - \tau(d^*))(\nu(d^*)\tau(d^*) + t_a t_u)\left(\nu''(d^*)(t_a - \tau(d^*)) + \tau''(d^*)(\nu(d^*) + t_u)\right) \\ &-t_u^2 \nu'(d^*)^2 (t_a - \tau(d^*))^3 - \tau'(d^*)^2(\nu(d^*) + t_u)^2 \left(\nu(d^*)(\nu(d^*)(t_a - \tau(d^*)) + 4t_c\tau(d^*)) + 4t_a t_u^2\right) \\ &+2t_u\nu(d^*)\nu'(d^*)\tau'(d^*)(t_a - \tau(d^*))^2(\nu(d^*) + t_u) \Big] \\ H_{den} = 64&(t_a - \tau(d^*))(t_u - \nu(d^*))^2(\nu(d^*)\tau(d^*) + t_a t_u)^2 \end{aligned}$$

Note that $H_{den} < 0$ as we have $(t_a - \tau(d^*)) < 0$ from Assumption 1. Rewriting H_{num} as

$$\begin{aligned}
H_{num} &= \tau(d^*)^2 \left[H1_{num}\left(H2_{num}\nu''(d^*) + H3_{num}\tau''(d^*)\right) + H4_{num} + H5_{num} + H6_{num}\right] \\
H1_{num} &= -4t_u^2(t_a - \tau(d^*))(\nu(d^*)\tau(d^*) + t_a t_u) > 0 \\
H2_{num} &= (t_a - \tau(d^*)) < 0 \\
H3_{num} &= (\nu(d^*) + t_u) > 0 \\
H4_{num} &= -t_u^2 \nu'(d^*)^2 (t_a - \tau(d^*))^3 \geq 0 \\
H5_{num} &= -\tau'(d^*)^2(\nu(d^*) + t_u)^2 \left(\nu(d^*)(\nu(d^*)(t_a - \tau(d^*)) + 4t_u\tau(d^*)) + 4t_a t_u^2\right) \lesseqgtr 0 \\
H6_{num} &= 2t_u\nu(d^*)\nu'(d^*)\tau'(d^*)(t_a - \tau(d^*))^2(\nu(d^*) + t_u) \leq 0
\end{aligned}$$

we can see that requiring $H_{num} < 0$ is equivalent to the condition

$$-\frac{1}{H1_{num}}\left(H4_{num} + H5_{num} + H6_{num}\right) > H2_{num}\nu''(d^*) + H3_{num}\tau''(d^*)$$

where $LHS \lesseqgtr 0$ while $RHS < 0$ due to our functional requirements on $\tau(\cdot)$ and $\nu(\cdot)$. The important thing to realize is that, firstly, the condition for negative definiteness reduces to a condition which is linear in $\nu''(d^*)$ and $\tau''(d^*)$, the curvature information of the targeting and the nuisance functions, and secondly, is given by an upper bound. If the sign of the upper bound is positive then this condition is always fulfilled as we have $RHS < 0$. Only if the sign of the upper bound is negative, the condition may bind. But then we can assume that $\tau(\cdot)$ is sufficiently concave and/or $\nu(\cdot)$ is sufficiently convex such that this condition holds since for our results we only require $\tau''(\cdot) < 0$ and $\nu''(\cdot) \geq 0$ which is in line with this condition.

A.3 Market Outcome

In equilibrium $p^* < 1$ and $\pi_i^*(a) \geq 0$. For this note that given equation (13), $p^* < 1$ if

$$2\frac{t_a t_u + \nu(d^*)\tau(d^*)}{\tau(d^*)t_u + \nu(d^*)\tau(d^*)} < 1 \iff t_a < \tau(d^*)\frac{(t_u - \nu(d^*))}{2t_u} < \tau(d^*) \tag{A.1}$$

By Assumption 1 we have that $\tau(d) > t_a$ for all d and therefore in particular also $\tau(d^*) > t_a$. Further, we have that $0 < (t_u - \nu(d^*))/2t_u < 1$, hence the last inequality. Thus, Assumption 1 is sufficient for the expression above to hold and $p^* < 1$.

Even the indifferent advertiser with highest transportation costs has positive profits in equilibrium because

$$\pi_i^*(\frac{1}{2}) = \frac{\tau(d^*)}{2} - \frac{t_a t_u + \nu(d^*)\tau(d^*)}{t_u + \nu(d^*)} - \frac{t_a}{2} \geq 0 \iff \tau(d^*)\frac{t_u - \nu(d^*)}{3t_u + \nu(d^*)} \geq t_a, \tag{A.2}$$

which is guaranteed by Assumption 1 for all d and especially for d^*. For this note that the term on the left in the last inequality is increasing in d.

A.4 Proofs of Propositions and Corollaries

Proof of Corollary 1

Proof. Rearranging terms in the first-order condition of platform profit maximization, given by equation (12), yields $2\kappa'(d^*) + \nu'(d^*) = \tau'(d^*)\frac{\nu(d^*)+t_u}{\tau(d^*)-t_a}$. By Assumption 1 we have $\tau(d^*) > t_a$. Hence the right hand side is positive, such that $2\kappa'(d^*)+\nu'(d^*) > 0$. □

Proof of Proposition 4

Proof. The proof relies on the monotonicity of the LHS and RHS in equations (7) and (12). Suppose, $\delta(d^*) > 1$ but $d^* < d^o$ and hence $\kappa'(d^*) < \kappa'(d^o)$. Using the implicit definition of d^o in (7) and d^* in (12) this implies $\delta(d^*)\tau'(d^*) - \nu'(d^*) < \tau'(d^o) - \nu'(d^o)$. Rearranging yields $\delta(d^*) < \frac{\tau'(d^o)}{\tau'(d^*)} + \frac{\nu'(d^*)-\nu'(d^o)}{\tau'(d^*)}$. But due to the curvature of $\tau(\cdot), \nu(\cdot)$ we have $\frac{\tau'(d^o)}{\tau'(d^*)} < 1$ and $\frac{\nu'(d^*)-\nu'(d^o)}{\tau'(d^*)} \leq 0$ for $d^* < d^o$, contradicting $\delta(d^*) > 1$. Now suppose $\delta(d^*) > 1$ and $d^* > d^o$, and hence $\delta(d^*) < \frac{\tau'(d^o)}{\tau'(d^*)} + \frac{\nu'(d^*)-\nu'(d^o)}{\tau'(d^*)}$. For $d^* > d^o$ we then have $\frac{\tau'(d^o)}{\tau'(d^*)} > 1$ and $\frac{\nu'(d^*)-\nu'(d^o)}{\tau'(d^*)} \geq 0$ and hence $\delta(d^*) > 1$. □

Proof of Proposition 7

Proof. To see that positive user prices in the two-sided model correspond to data overprovision in the one-sided pricing model, note that user prices are positive in the two-sided pricing model if $\tau(d^o) - \nu(d^o) < t_a + t_u$. From Proposition 4 we know that in the one-sided pricing model too little data is provided if $t_a + t_u < \tau(d^*) - \nu(d^*)$. But this would mean that $d^* < d^o$, which contradicts $\tau(d^o) - \nu(d^o) < t_a + t_u < \tau(d^*) - \nu(d^*)$, as $\tau(d)$ is increasing and $\nu(d)$ decreasing in d. Hence it can only be that in the one-sided model there is overprovision, such that $d^* > d^o$ and $\tau(d^o) - \nu(d^o) < \tau(d^*) - \nu(d^*) < t_a + t_u$.
To see that negative user prices in the two-sided model correspond to data underprovision in the one-sided pricing model, note that user prices are negative in the two-sided pricing model if $\tau(d^o) - \nu(d^o) > t_a + t_u$. From Proposition 4 we know that too much data is provided if $t_a + t_u > \tau(d^*) - \nu(d^*)$. But this would mean that $d^* > d^o$, which contradicts $\tau(d^o) - \nu(d^o) > t_a + t_u > \tau(d^*) - \nu(d^*)$. Hence it must be that in the one-sided model there is underprovision, such that $d^* < d^o$ and $\tau(d^o) - \nu(d^o) > \tau(d^*) - \nu(d^*) > t_a + t_u$. □

A.5 Proofs for Comparative Statics

For the effect of transportation costs t_a and t_u on d_i note first that

$$\left.\frac{\partial X_i}{\partial d_i}\right|_{\substack{d_i=d^*\\p_i=p^*}} = -\frac{\{t_a t_u + \nu(d^*)\left[t_a p^* + (1-p^*)\tau(d^*)\right]\}\,\tau'(d^*)}{4\left[\tau(d^*)-t_a\right]\left[t_a t_u + (1-p^*)\nu(d^*)\tau(d^*)\right]} < 0, \tag{A.3}$$

$$\left.\frac{\partial A_i}{\partial d_i}\right|_{\substack{d_i=d^*\\p_i=p^*}} = -\frac{(1-p^*)\left[t_a t_u + \nu(d^*)\tau(d^*)\right]\tau'(d^*)}{4\left[\tau(d^*)-t_a\right]\left[t_a t_u + (1-p^*)\nu(d^*)\tau(d^*)\right]} < 0, \tag{A.4}$$

because $\tau'(d^*) > 0$, while $\tau(d^*) > t_a$ by Assumption 1 and $p^* < 1$ as established in Section A.3. Differentiating (A.3) and (A.4) with respect to transportation costs yields

$$\left.\frac{\partial^2 X_i}{\partial d_i \partial t_a}\right|_{\substack{d_i=d^*\\p_i=p^*}} = -\frac{(1-p^*)\nu(d^*)\left[t_a t_u + \nu(d^*)\tau(d^*)\right]\tau'(d^*)}{4\left[\tau(d^*)-t_a\right]\left[t_a t_u + (1-p^*)\nu(d^*)\tau(d^*)\right]^2} < 0,$$

$$\left.\frac{\partial^2 A_i}{\partial d_i \partial t_a}\right|_{\substack{d_i=d^*\\p_i=p^*}} = \frac{(1-p^*)t_u\left[t_a t_u + \nu(d^*)\tau(d^*)\right]\tau'(d^*)}{4\left[\tau(d^*)-t_a\right]\left[t_a t_u + (1-p^*)\nu(d^*)\tau(d^*)\right]} > 0,$$

$$\left.\frac{\partial^2 X_i}{\partial d_i \partial t_u}\right|_{\substack{d_i=d^*\\p_i=p^*}} = \frac{t_a\,\{t_a t_u + \nu(d^*)\left[t_a p^* + (1-p^*)\tau(d^*)\right]\}\,\tau'(d^*)}{4\left[\tau(d^*)-t_a\right]\left[t_a t_u + (1-p^*)\nu(d^*)\tau(d^*)\right]} > 0,$$

$$\left.\frac{\partial^2 A_i}{\partial d_i \partial t_u}\right|_{\substack{d_i=d^*\\p_i=p^*}} = \frac{(1-p^*)\tau(d^*)\,\{t_a t_u + \nu(d^*)\left[t_a p^* + (1-p^*)\tau(d^*)\right]\}\,\tau'(d^*)}{4\left[\tau(d^*)-t_a\right]\left[t_a t_u + (1-p^*)\nu(d^*)\tau(d^*)\right]} > 0.$$

Further note that

$$\left.\frac{\partial^2 A_i}{\partial d_i \partial t_a} - \frac{\partial^2 X_i}{\partial d_i \partial t_a}\right|_{\substack{d_i=d^*\\p_i=p^*}} = \frac{(1-p^*)\left[t_u + \nu(d^*)\right]\left[t_a t_u + \nu(d^*)\right]\tau'(d^*)}{4\left[\tau(d^*)-t_a\right]\left[t_a t_u + (1-p^*)\nu(d^*)\tau(d^*)\right]} > 0.$$

To see that $\mathrm{d}\Pi_i^P/\mathrm{d}t_u < 0$, note that

$$\begin{aligned}
\frac{\mathrm{d}\Pi_i^P}{\mathrm{d}t_u} &= \frac{-\left[\tau(d^*)-t_a\right]\left[\nu(d^*) - t_u\nu'(d^*)\frac{\mathrm{d}d^*}{\mathrm{d}t_u}\right] + \frac{\mathrm{d}d^*}{\mathrm{d}t_u}\nu(d^*)\tau'(d^*)\left[t_u+\nu(d^*)\right]}{\left[t_c+\nu(d^*)\right]^2}\\
&= \frac{\left[\tau(d^*)-t_a\right]^2\left[\left(t_u+\nu(d^*)\right)\nu'(d^*)\tau'(d^*) - \nu(d^*)\left\{\left(\tau(d^*)-t_a\right)\left[\kappa''(d)+\nu''(d^*)\right] - t_c\left(t_u+\nu(d^*)\right)\tau''(d^*)\right\}\right]}{\left[t_u+\nu(d^*)\right]^2\Psi(d^*)}\\
&< 0,
\end{aligned} \tag{A.5}$$

where $\mathrm{d}d^*/\mathrm{d}t_u$ is from equation (15), while the term in the denominator is given by

$$\begin{aligned}
\Psi(d^*) = {} & \left[2\kappa''(d^*) + \nu''(d^*)\right]\left(\tau(d^*)-t_a\right)^2 - \nu'(d^*)\tau'(d^*)\left(\tau(d^*)-t_a\right)\\
& + \left(\nu(d^*)+t_u\right)\left[\tau'(d^*)^2 - \left(\tau(d^*)-t_a\right)\tau''(d^*)\right] > 0.
\end{aligned} \tag{A.6}$$

Note for the inequalities that $\tau'(d^*) > 0$, $\tau''(d^*) < 0$ while $\nu'(d^*) \leq 0$, $\nu''(d^*) \geq 0$ by construction, and $\tau(d^*) > t_a$ by Assumption 1.

B Online Appendix

In this online appendix we will provide derivations and intuition for comparative statics not covered in the main text. Further, we present extensions to our baseline model where we consider multi-homing, elastic total demand and positive cross-group externalities.

B.1 Comparative Static Effects on Prices in Equilibrium

For this analysis consider the platform's first-order condition in equation (11) and note that the price depends indirectly on the effects of p_i on advertiser and user market shares A_i and X_i as given in Section 4.1.

$$\left.\frac{\partial X_i}{\partial p_i}\right|_{\substack{d_i=d^*\\p_i=p^*}} = \frac{\nu(d^*)\tau(d^*)}{4\left[t_a t_u + (1-p^*)\nu(d^*)\tau(d^*)\right]} > 0, \tag{B.1}$$

$$\left.\frac{\partial A_i}{\partial p_i}\right|_{\substack{d_i=d^*\\p_i=p^*}} = -\frac{t_u\tau(d^*)}{4\left[t_a t_u + (1-p^*)\nu(d^*)\tau(d^*)\right]} < 0, \tag{B.2}$$

because $p^* < 1$ as established in Appendix A.

Competition for Advertisers

Note that in Section 5 we discussed that lower advertiser-side competition intensity increases the level of data collection in equilibrium, i.e. $\mathrm{d}d^*/\mathrm{d}t_a > 0$. Here we analyze the effects of competition intensity for advertisers on p^*. Differentiating (B.1) with respect to transportation costs t_a yields

$$\left.\frac{\partial^2 X_i}{\partial p_i \partial t_a}\right|_{\substack{d_i=d^*\\p_i=p^*}} = -\frac{t_u\tau(d^*)\nu(d^*)}{4\left[t_a t_u + (1-p^*)\nu(d^*)\tau(d^*)\right]^2} < 0,$$

$$\left.\frac{\partial^2 A_i}{\partial p_i \partial t_a}\right|_{\substack{d_i=d^*\\p_i=p^*}} = \frac{t_u^2\,\tau(d^*)}{4\left[t_a t_u + (1-p^*)\nu(d^*)\tau(d^*)\right]^2} > 0.$$

Further note that

$$\frac{\partial^2 A_i}{\partial p_i \partial t_a} - \frac{\partial^2 X_i}{\partial p_i \partial t_a}\Bigg|_{\substack{d_i=d^* \\ p_i=p^*}} = \frac{t_u \left[t_u + \nu(d^*)\right] \tau(d^*)}{4 \left[t_a t_u + (1-p^*)\nu(d^*)\tau(d^*)\right]^2} > 0.$$

If competition for advertisers softens, i.e. transportation costs t_a increase, advertisers become less sensitive to changes in prices such that $\partial^2 A_i/(\partial p_i \partial t_a) > 0$. Consequently, users become more sensitive to prices (which repel advertisers) such that $\partial^2 X_i/(\partial p_i \partial t_a) < 0$. Overall, the former effect dominates the latter effect in magnitude. Consequently, and as $X_i^* = A_i^* = 1/2$, the right-hand-side of equation (11) increases in t_a such that the equilibrium price must rise, i.e.

$$\frac{\mathrm{d}p^*}{\mathrm{d}t_a} > 0. \tag{B.3}$$

Intuitively, higher advertiser transportation costs mean more sticky advertisers and hence decreased platform competition for advertisers. Therefore, it is straightforward that advertiser prices rise, which is line with standard intuition.

Competition for Users

In Section 5 we discussed that lower competition intensity for users decreases platforms' equilibrium level of data collection, i.e. $\mathrm{d}d^*/\mathrm{d}t_u > 0$. Here we analyze the effects of competition intensity for users on p^*. Differentiating (B.1) with respect to transportation costs t_u yields

$$\frac{\partial^2 X_i}{\partial p_i \partial t_u}\Bigg|_{\substack{d_i=d^* \\ p_i=p^*}} = -\frac{t_a \tau(d^*)\nu(d^*)}{4 \left[t_a t_u + (1-p^*)\nu(d^*)\tau(d^*)\right]^2} < 0,$$

$$\frac{\partial^2 A_i}{\partial p_i \partial t_u}\Bigg|_{\substack{d_i=d^* \\ p_i=p^*}} = -\frac{(1-p^*)\nu(d^*)\tau(d^*)^2}{4 \left[t_a t_u + (1-p^*)\nu(d^*)\tau(d^*)\right]^2} < 0.$$

If competition for users softens, i.e. transportation costs t_u increase, users become less sensitive to changes in prices such that $\partial^2 X_i/(\partial p_i \partial t_u) < 0$. Consequently, advertisers, too, become less sensitive to prices (which now repel less users) such that $\partial^2 X_i/(\partial d_i \partial t_u) < 0$. Therefore the right-hand-side of equation (11) decreases in t_u such that the equilibrium price must fall, i.e.

$$\frac{\mathrm{d}p^*}{\mathrm{d}t_u} < 0. \tag{B.4}$$

Again, this is in line with standard platform intuition: advertiser prices fall if the user side becomes less sensitive (elastic).

Nuisance

First, we consider the effects of nuisance on data collection.[20] Totally differentiating the first-order conditions from equations (12) and (13) w.r.t. $\nu(d)$ and solving for $\mathrm{d}d^*/\mathrm{d}\nu(d)$ yields

$$\left.\frac{\mathrm{d}d^*}{\mathrm{d}\nu(d)}\right|_{d=d^*} = \frac{(\tau(d^*) - t_a)\,\tau'(d^*)}{\Psi(d^*)} > 0. \tag{B.5}$$

Second, we evaluate the effects of nuisance on p^*. Solving for $\mathrm{d}p^*/\mathrm{d}\nu(d)$ and dropping the argument d^* of $\nu(d^*)$ and $\tau(d^*)$ to abbreviate, yields

$$\left.\frac{\mathrm{d}p^*}{\mathrm{d}\nu(d)}\right|_{d=d^*} = \frac{-2t_u\,(t_a-\tau)^2\left[\tau\left(\tau''\,(\nu+t_u) - (\tau-t_a)\,(2\kappa''+\nu'')\right) - (\nu+t_u)\,\tau'^2\right]}{(\nu+t_u)\,\tau^2\,\Psi(d^*)} > 0, \tag{B.6}$$

where $\Psi(d^*)$ is defined in equation (A.6). Intuitively, higher (absolute) nuisance results in lower user demand. To counterbalance this effect, platforms would increase ad prices as ads become relatively less attractive. Additionally, more user data would be collected in order to soften the nuisance increase. Interpreted from the point of view of users, they are now willing to incur marginally more privacy costs in order to obtain some nuisance reduction.

Targeting

First, we consider the effects of the targeting technology on data collection.[21] Solving for $\mathrm{d}d^*/\mathrm{d}\tau(d)$ yields

[20] Note that nuisance is a function in our model. To assess an increase in nuisance we treat it as fixed and consider an upward shift, without changing any curvature. For this we slightly abuse notation to stay consistent with the rest of our comparative statics, such that e.g. by $\mathrm{d}d^*/\mathrm{d}\nu(d)|_{d=d^*}$ we intuitively consider the effect of adding a positive constant c to the function, i.e. $\nu(d)+c$ where $c>0$, on d^*.

[21] Note that targeting is a function, which we treat as fixed here, such that comparative statics are performed as described in footnote 20.

$$\left.\frac{\mathrm{d}d^*}{\mathrm{d}\tau(d)}\right|_{d=d^*} = -\frac{(\nu(d^*)+t_u)\,\tau'(d^*)}{\Psi(d^*)} < 0. \tag{B.7}$$

Second, we evaluate the effects of nuisance on p^*. Solving for $\mathrm{d}p^*/\mathrm{d}\tau(d)$ and dropping again the argument d^* to abbreviate, yields

$$\left.\frac{\mathrm{d}p^*}{\mathrm{d}\tau(d)}\right|_{d=d^*} = \frac{2t_u\,(\tau - t_a)\,[\tau''\,(\nu+t_u)\,t_a - (\tau - t_a)\,[\nu'\tau' + t_a\,(2\kappa''+\nu'')]]}{(\nu+t_u)\,\tau^2\Psi(d^*)} \gtrless 0. \tag{B.8}$$

platforms to create the same ad value with less personal data, hence in equilibrium platforms will compete to 'relax' the data requirement for users. Two effects are relevant for the effect on ad prices. On the one hand ads become more valuable, hence platforms might increase the price, i.e. their share, of this value (intensive margin). On the other hand, platforms might prefer to attract more of these valuable advertisers by reducing the ad price (extensive margin). Overall, the effect on ad prices depends on which of the opposing effects is stronger.

B.2 Comparative Static Effects on Platform Profits, Advertiser Profits and User Utility

In this section we provide further intuition on equilibrium profits and utility by presenting comparative statics.

Effects on Platform Profits

The effects on platform profits $\Pi_i^* = p^*\tau(d^*)\,X_i^*\,A_i^* = (1/4)\,p^*\tau(d^*)$ can be written as

$$\frac{\mathrm{d}\Pi_i^*}{\mathrm{d}z} = \frac{1}{4}\left[\frac{\mathrm{d}p^*}{\mathrm{d}z}\tau(d^*) + \tau'(d^*)\frac{\mathrm{d}d^*}{\mathrm{d}z}p^*\right]. \tag{B.9}$$

We look at the effects of advertiser competition intensity. For $z = t_a$ both terms on the right-hand side are positive and hence $\mathrm{d}\Pi_i^*/\mathrm{d}t_a > 0$. Intuitively, when competition for advertisers becomes more intense (t_a decreases), then prices for ad-placing decrease. In turn, less data is collected from users, such that targeting becomes less effective, and less total revenue is made on the ad market. Both these effects decrease platform profits.

The intensity of user-side competition increases platforms' surplus, i.e. $\mathrm{d}\Pi_i^*/\mathrm{d}t_u < 0$. This effect is discussed in the main text in Section 5.

Increased nuisance (higher $z = \nu(d)$) increases platforms' surplus, i.e. $\mathrm{d}\Pi_i^*/\mathrm{d}\nu(d) > 0$. More data is collected, which increases targeting and hence the (residual) value of a placed ad, thus also higher prices can be sustained. Overall, this unambiguously benefits platforms.

Increased targeting (higher $z = \tau(d)$) increases platforms' surplus, i.e. $\mathrm{d}\Pi_i^*/\mathrm{d}\tau(d) > 0$. Although less data is collected, the absolute externality of users, i.e. targeting, increases the value to be shared between platforms and advertisers. While the effect on prices remains ambiguous, overall, platforms benefit. To see that note that

$$\frac{\mathrm{d}\Pi_i^*}{\mathrm{d}\tau(d)} = \frac{\tau(d^*) - t_a}{(t_u + \nu(d^*))\,\Psi(d^*)} \left[- (t_u + \nu(d^*))\,\nu'(d^*)\tau'(d^*) \right.$$
$$\left. + \nu(d^*) \left\{ (\tau(d^*) - t_a) \left[\kappa''(d) + \nu''(d^*) \right] - t_c \left(t_u + \nu(d^*) \right) \tau''(d^*) \right\} \right] > 0, \tag{B.10}$$

where $\mathrm{d}d^*/\mathrm{d}t_u$ is from equation (15), while $\Psi(d^*)$ is defined in equation (A.6).

Effects on Advertiser Profits

The effects on advertiser profits $\pi_i^*(a) = (1-p^*)\tau(d^*)X_i^* - t_a|l_i - a| = (1/2)\,(1-p^*)\,\tau(d^*) - t_a|l_i - a|$ are given by

$$\frac{\mathrm{d}\pi_i^*(a)}{\mathrm{d}z} = \frac{1}{2} \left[-\frac{\mathrm{d}p^*}{\mathrm{d}z} \tau(d^*) + \tau'(d^*) \frac{\mathrm{d}d^*}{\mathrm{d}z} (1 - p^*) \right] - |l_i - a| \frac{\mathrm{d}t_a}{\mathrm{d}z}. \tag{B.11}$$

Stronger competition for advertisers (lower $z = t_a$) makes advertisers overall better off, i.e. $\mathrm{d}\pi_i^*/\mathrm{d}t_a < 0$. However, there are multiple effects at work. Firstly, prices fall, such that the first term on the right hand side increases. Secondly, less personal data from users can be collected, which makes targeting less effective, therefore the second term is negative. Thirdly, also transportation costs decrease, which increases advertiser profits. Overall, the price and transportation cost reduction effects outweigh decreased targeting effectiveness. For this note that

$$\frac{\mathrm{d}\pi^A}{\mathrm{d}t_a} = \frac{1}{4\left[t_u + \nu(d^*)\right]^2} \left\{ -6t_u\nu(d^*) - \nu(d^*)^2 \left[1 + 2\tau'(d^*) \frac{\mathrm{d}d^*}{\mathrm{d}t_a} \right] \right.$$

$$
\begin{aligned}
& + t_c \left[-4\nu'(d^*) \frac{\mathrm{d}d^*}{\mathrm{d}t_a} \left(\tau(d^*) - t_a\right) + t_c \left(-5 + 2\tau'(d^*) \frac{\mathrm{d}d^*}{\mathrm{d}t_a} \right) \right] \Bigg\} \\
= & -\frac{1}{4\left[t_u + \nu(d^*)\right] \Psi(d^*)} \Bigg\{ -\nu'(d^*) \left(t_u + \nu(d^*)\right) \left(\tau(d^*) - t_a\right) \tau'(d^*) + 3 \left(t_u + \nu(d^*)\right)^2 \tau'(d^*)^2 \\
& - \left(5t_u + \nu(d^*)\right) \left(\tau(d^*) - t_a\right) \left[-\nu''(d^*) \left(\tau(d^*) - t_a\right) + \left(t_u + \nu(d^*)\right) \tau''(d^*) \right] \Bigg\} \\
< & \ 0,
\end{aligned}
\tag{B.12}
$$

where $\mathrm{d}d^*/\mathrm{d}t_a$ is from equation (14), while $\Psi(d^*)$ is defined in equation (A.6).

Stronger competition for users (increase $z = t_u$) hurts advertisers, hence $\mathrm{d}\pi_i^A/\mathrm{d}t_u > 0$. The platforms' bottleneck position allows them to increase prices (negative first term) and, further, less user data can be collected, such that targeting becomes less effective (negative second term).

Increased nuisance (higher $z = \nu(d)$) decreases advertisers' surplus, i.e. $\mathrm{d}\pi_i^A/\mathrm{d}\nu(d) < 0$. Although more data is collected, which increases targeting and hence the value of a placed ad, also prices increase. Overall, this hurts advertisers. To see that, note

$$
\begin{aligned}
\frac{\mathrm{d}\pi^A}{\mathrm{d}\nu(d)} = & -\frac{\left[\tau(d^*) - t_a\right]}{2\left[t_u + \nu(d^*)\right]^2 \Psi(d^*)} \Bigg\{ \left(t_u + \nu(d^*)\right)^2 \tau'(d^*)^2 \\
& + 2\left[\tau(d^*) - t_a\right] t_c \left\{ \left(\tau(d^*) - t_a\right) \left[\kappa''(d) + \nu''(d^*)\right] - t_c \left(t_u + \nu(d^*)\right) \tau''(d^*) \right\} \Bigg\} \\
< & \ 0,
\end{aligned}
\tag{B.13}
$$

Increased targeting (higher $z = \tau(d)$) has an ambiguous effect on advertisers' surplus. While the targeting function becomes better, less data needs be collected which again reduces targeting effectiveness. Further, the effect on prices is ambiguous. Hence, overall effects on advertiser surplus remain unclear.

Effects on User Utility

The effects on a user's utility $u_i^*(x) = \underline{u} - \kappa(d^*) - \nu(d^*)A_i^* - t_u|l_i - x| = \underline{u} - \kappa(d^*) - (1/2)\,\nu(d^*) - t_u|l_i - x|$ are given by

$$
\frac{u_i^*(x)}{\mathrm{d}z} = -\frac{\mathrm{d}d^*}{\mathrm{d}z} \left[\kappa'(d^*) + \frac{\nu'(d^*)}{2} \right] - \frac{\mathrm{d}t_u}{\mathrm{d}z} |l_i - x|. \tag{B.14}
$$

Note that by Corollary 1 the term in brackets on the right-hand side is positive and that for $z \in \{t_a, t_u\}$ we have $\mathrm{d}d^*/\mathrm{d}z > 0$ such that $\mathrm{d}u_i/\mathrm{d}z < 0$.

Intuitively, less competition for advertisers (higher $z = t_a$) increases the amount of data collected in equilibrium, which overall leaves users worse off, as privacy concerns are increased, although ads are more targeted and hence nuisance smaller.

Less competition for users (higher $z = t_u$) increases the amount of data collected, such that privacy concerns are increased, although it reduces nuisance costs. Further strengthened by increased transportation costs for users, quite naturally users' utility overall decreases.

Increased nuisance (higher $z = \nu(d)$) decreases users' utility, i.e. $\mathrm{d}u_i/\mathrm{d}\nu(d) < 0$ because again more data is collected.

Increased targeting (higher $z = \tau(d)$) increases users' utility, i.e. $\mathrm{d}u_i/\mathrm{d}\tau(d) < 0$. Although targeting does not directly affect users, less data is collected, which is beneficial for users.

B.3 Market Coverage and Multi-homing

Advertiser Side

We start this section by lifting Assumption 1 for full market coverage and the single-homing assumption for advertisers. Analytically, this is achieved by pinning down advertisers which are indifferent between joining a platform and abstaining such that the total mass of advertisers joining platform i is determined by $\pi_i(a) = 0$.

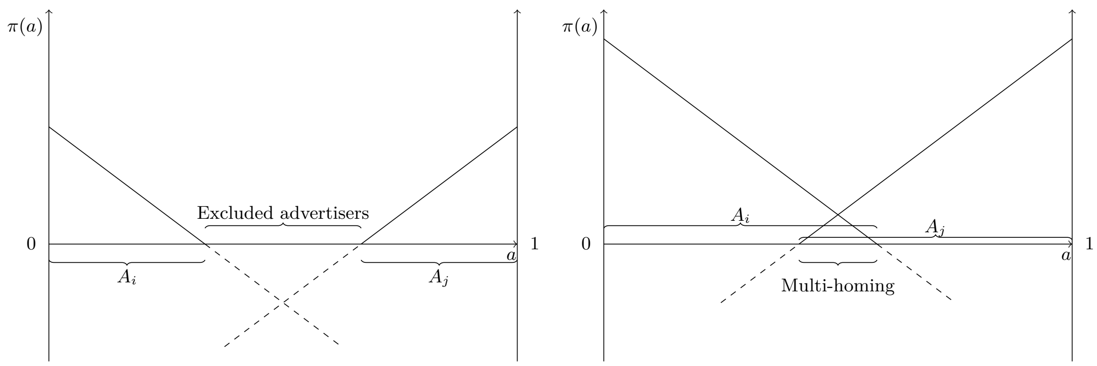


FIGURE B.1: Relaxed Advertiser Market Assumption

Figure B.1 shows two potential outcomes of this alternative setup. In the first case the total mass of participating advertisers in the market is smaller than 1 while advertisers 'in the middle' choose not to participate as their transportation costs are too high. In the second case the sets of advertisers joining platform i and j are overlapping such that advertisers 'in the middle' join both platforms, i.e. they multi-home. The remaining analysis follows the steps from the baseline model and is omitted at this point.

The welfare maximizing level of data d_a^o is then given by

$$\kappa'(d_a^o) = A_i^o(d_a^o)\tau'(d_a^o) - A_i^o(d_a^o)\nu'(d_a^o) \tag{B.15}$$

where $A_i^o(d)$ denotes the symmetric mass of advertisers on each platform and is given by $A_i^o(d) = [\tau(d) - \nu(d)]/(2t_a)$. The equilibrium level of data under platform competition d_a^* is then given by

$$\kappa'(d_a^*) = \left(A_i^*(d_a^*)\frac{\nu(d_a^*)}{\tau(d_a^*)} + \frac{t_u}{\tau(d_a^*)}\right)\tau'(d_a^*) - A_i^*(d_a^*)\nu'(d_a^*) \tag{B.16}$$

while $A_i^*(d_a^*) = [(1 - p_a^*(d_a^*))\tau(d_a^*)]/(2t_a)$. We can see immediately that whether the resulting allocation is an equilibrium with multi-homing or with excluded advertisers depends on functional forms and parameters. We will therefore discuss the two cases separately in the following.

Assume transport costs t_a are sufficiently low to allow a multi-homing allocation of advertisers under the efficient benchmark, i.e. $A_i^o(d_a^o) > 1/2$. Comparing the condition for the resulting efficient level of data provision to our baseline condition in (12) we see that $d_a^o > d^o$, under multi-homing the efficient level of data provision is higher than under single-homing. The idea is that additional advertisers are attracted in order to maximize total value creation in the economy. The comparison of the new competitive level of data provision d_a^* to the new efficiency benchmark as well as to our baseline model is, however, ambiguous. As competition for advertisers is now relaxed, platforms might not be forced to offer high levels of d to attract additional advertisers. At the same the value creation aspect from a larger total number of advertisers is still valid, such that the net effect on the level of data provision remains ambiguous.

When transportation costs t_a are sufficiently large, some advertisers 'in the middle' would not join any platform, such that $A_i^o(d_a^o) < 1/2$ and also $A_i^*(d_a^*) < 1/2$. Note that the efficient level is then also lower than in our benchmark $d_a^o < d^o$ as attracting advertisers becomes relatively expensive and it becomes more efficient to exclude some advertisers than to offer very high levels of d. The comparison to the market outcome, however, remains ambiguous. While the same efficiency argument applies, platforms also have an additional incentive to increase their intensive margin by increasing d to offset the reduction in advertising demand. Again, depending on functional forms either effect may dominate.

User Side

Similarly on the user side, by relaxing Assumption 2 it is possible that $\underline{u}$ becomes sufficiently small relative to transportation costs, such that users 'in the middle' prefer to abstain from both platforms. If $\underline{u}$ is sufficiently large relative to transportation costs, users 'in the middle' might choose to join both platforms. In both cases user market shares are determined through the utility of the indifferent user relative to the outside option.

The symmetric welfare-maximizing level of data d_u^o is then given by

$$\kappa'(d_u^o) = X_i^o(d_u^o)\frac{t_u}{\tau(d_u^o) + 2\underline{u} - 2\kappa(d_u^o) - \nu(d_u^o)}\tau'(d_u^o) - \frac{1}{2}\nu'(d_u^o), \tag{B.17}$$

where $X_i^o(d_u^o)$ denotes the symmetric mass of users on each platform and is given by $X_i^o(d_u^o) = [2\underline{u} - 2\kappa(d_u^o) - \nu(d_u^o)]/(2t_u)$. The equilibrium level of data under platform competition d_u^* is then given by

$$\kappa'(d_u^*) = X_i^*(d_u^*)\frac{t_u}{\tau(d_u^*)}\tau'(d_u^*) - \frac{1}{2}\nu'(d_u^*), \tag{B.18}$$

while $X_i^*(d_u^*) = [2\underline{u} - 2\kappa(d_u^*) - \nu(d_u^*)]/(2t_u)$. From this we can immediately see that $d_u^* > d_u^o$, i.e. there is always over-provision of user data. While the efficient benchmark takes into account the tradeoff between excluding users and total value creation, the market outcome only compares the targeting benefit to the exclusion of users. Further note that if the market is not covered such that $X_i(d_u) < 1/2$, the efficient as well as the equilibrium level of data provision is lower than in the baseline model, i.e. $d_u^o < d^o$ and $d_u^* < d^*$ because $t_u/\tau(d) < \delta(d)\ \forall d$.

It is worthwhile to note that under user multi-homing as well as under relaxed user market coverage we get that $\mathrm{d}d_u^*/\mathrm{d}t_u = \mathrm{d}d_u^*/\mathrm{d}t_a = 0$, i.e. the transportation cost parameters on either market side are irrelevant for the equilibrium (and also the efficient) level of data collection. This is because t_u now merely scales demand while the relevant trade-off for the choice of d involves the actual utility from joining the platform, which is unaffected by the demand scale.

Under this setup user demand becomes more elastic than in the baseline model which undermines platforms' incentive to increase d. At the same time platforms would also increase prices $\mathrm{d}p_u^*/\mathrm{d}t_u > 0$ if it becomes increasingly difficult to attract users. Note that we seemingly found the opposite effect in our baseline model $\mathrm{d}p^*/\mathrm{d}t_u < 0$, however, the interpretation of t_u changes substantially such that the two results do not contradict each other: the harder it is to keep users, the higher the prices for advertisers.

In fact platforms are able to overcompensate the reduction in user demand such that $\mathrm{d}\Pi_u^*/\mathrm{d}t_u > 0$ (and for advertisers $\mathrm{d}\pi_u^*/\mathrm{d}t_u < 0$). Again, as the interpretation of t_u essentially reverses, we had the opposite results in our baseline model where platform profits decreased in t_u (while advertiser profits increased). This is also reflected in the effect on the advertiser side where equilibrium prices rise in t_a under both model specifications, i.e. $\mathrm{d}p_u^*/\mathrm{d}t_a > 0$ as the interpretation remains identical.

B.4 Positive Cross-group Externalities

Consider the following modification of the users' utility function:

$$u_i(x) = \underline{u} - \kappa(d_i) + \rho(d_i)A_i - t_u|l_i - x|. \tag{B.19}$$

The concave and twice-differentiable function $\rho(d)$ represents the relevance from a user's point of view of seeing A_i offers, where $\rho'(d) \geq 0$ and $\rho''(d) \leq 0$. However, $\rho(d)$ can now be entirely negative, positive or might even switch signs. The first case is discussed in depth in the main paper, where we consider the case $\rho(d) = -\nu(d)$. The second case, a strictly positive effect, can be thought of as a traditional 'dating' model, where one group strictly enjoys the presence of the other group. The last case can be thought of as a more nuanced version of our nuisance cost in the baseline model. While for low values of d, i.e. the platform has very little information about the consumer, a user dislikes the interaction with the other market side, the interaction might turn out to be valuable once the platform has sufficient information, i.e. d is sufficiently large. A typical example would be the recommendation system on Amazon. While it is debatable, whether Amazon is a two-sided market in the traditional sense, the product recommendation system might serve as a useful example. A new customer might see all kind of product recommendations, some of which are completely useless to the user and are just a waste of attention. However, once Amazon has acquired sufficient information about the user's preferences through analyzing the purchasing and browsing history, the recommendations become more personalized, and the user finds actual value in looking through them.

From a modelling perspective we only require that the relevance is monotonically increasing in the amount of data, but with decreasing returns. Since the curvature of the maximization problem therefore remains unchanged, the characterization of the second order conditions given in the Appendix A also remain qualitatively unchanged. The absolute value of the function $\rho(d)$ is in the end of minor importance regarding the key mechanics of the model, however, it has to be taken care of through appropriately adjusting the modelling assumptions. In order to assure full market coverage on the offer side, we now have the following set of assumptions.

Assumption 3. *Competition for advertisers is sufficiently strong, i.e.* $t_a \leq \bar{t}_a$*. For this, it is necessary that competition for users is sufficiently weak and that there are gains of trade for all advertisers, even without data collection, i.e.*

$$\begin{aligned} &(a) \quad t_u > |\rho(0)|, \rho(d) < t_u \\ &(b) \quad t_a < \tau(0) \end{aligned}$$

The upper bound on t_a is then given by $\bar{t}_a := \frac{t_u\tau(0)+\rho(0)\tau(0)}{3t_u-\rho(0)}$. Since now net cross-group externalities might be positive, a problem of platform tipping must be taken into account. In particular the following assumption ensures that the competitive symmetric equilibrium leads to positive prices (and therefore positive platform profits), so that a platform would not be indifferent whether to enter the market if just one platform serves the entire market.

Assumption 4. *To ensure market participation of both platforms it is necessary to have*

$$t_a t_u > \rho(\cdot)\tau(\cdot).$$

Note that for negative $\rho(\cdot)$ as in our main model, this assumption is always fulfilled as then the RHS is always negative, while the LHS is always positive. Accordingly, if $\rho(\cdot)$ switches signs, the range in which $\rho(\cdot)$ is negative is unproblematic. Therefore the only potentially problematic case is if $\rho(\cdot)$ is positive or can turn positive since it further restricts the parameter space in addition to the previous assumption.[22] Given that both assumptions are satisfied, the analysis is analogous to our main model and all major results still hold.

[22] In the following we sketch a set of conditions under which both assumptions would be satisfied. Note Assumption 4 specifies a lower bound $t_a > \underline{t_a}$ with $\underline{t_a} \equiv \frac{1}{t_u}\rho(\cdot)\tau(\cdot)$. It is therefore necessary to show that the set of t_a satisfying Assumptions 3 and 4 is non-empty. In particular, if it holds that $lim_{d\to\infty}\underline{t_a} < \bar{t}_a$ we can always find intermediate values of t_a satisfying both conditions. For this to be the case it is necessary that $lim_{d\to\infty}\underline{t_a} < \tau(0)$ and that $\rho(\cdot)$ is small if positive.

Chapter III

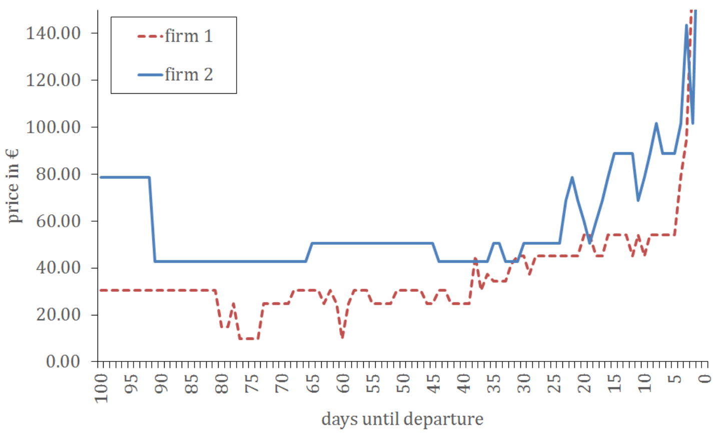
Dynamic Pricing under Capacity-constrained Competition with a Deadline

1 Introduction

In many markets we observe volatile price paths throughout the entire selling time, for example in markets for travel or event bookings. Many of these markets share the following characteristics. First, there is a fixed deadline after which the goods expire or perish. An airline ticket for an already departed flight or a hotel booking for yesterday night is worthless today. Second, firms hold only limited capacities. Hotels have a fixed number of rooms and airlines have to predetermine aircraft types with fixed numbers of seats to serve a certain route on a certain day. Third, firms can engage in dynamic pricing behavior, i.e. they are able to adjust prices throughout the entire selling horizon. And fourth, consumers might act in a forward-looking manner, allowing them to strategically delay their purchase in the hope of better prices. Examples for these markets are ample. Besides classic travel (airline, train, bus), hotel or event (theater, sports, concert) booking markets, one could also think of seasonal goods markets, such as markets for Christmas trees or first-day-of-school items. Additionally, many of the above named goods are also traded on secondary markets with similar characteristics.

As a real-world example path of dynamic prices consider Figure 1. Prices posted by two competing airlines for flights on the same route at the same time of the day are depicted for 100 days prior to the departure day.

Figure 1: Observed Airline Price Path



Based on self-collected price data of Ryanair (firm 1) and Aegean Air (firm 2), competing on the ATH-SKG morning route of May 30, 2014.

In this paper I construct a simple dynamic competition model of discrete and finite periods, which provides an explanation for these volatile price paths we experience in many real markets with the above characteristics. Oligopolists with ex-ante determined capacities post prices in each time period up to the deadline (sometimes referred to as closed loop-pricing). In each such period one consumer arrives and decides if and where to purchase. While firms' goods provide a basic value, there is additional consumer heterogeneity. Employing a discrete choice model with idiosyncratic taste shocks (multinomial logit), the resulting demand structure yields choice probabilities for each firm's good and also a choice probability for not buying (possibly waiting). In the airline or hotel application this heterogeneity could be due to differences in brand taste or variation in loyalty program membership across consumers.

Each period can be characterized by its state, consisting of information on the number of remaining selling periods as well as firms' remaining capacity stocks and the number of consumers possibly waiting in the market. Firms know the state and anticipate that there will be a state transition into the next period, depending on the specifics of possible trade.[1] Using inter-temporal value functions for firms we get a dynamic programming framework. There exists a pure-strategy price equilibrium in each period, where firms consider current payoffs as well as possible implications for future states and payoffs, giving rise to a Sub-game Perfect Nash Equilibrium for the whole game. Furthermore, I

[1]While myopic consumers choosing between buying or exiting the market do not need knowledge of the state, forward-looking consumers can choose to wait in the market and hence need knowledge of the state and its transition.

provide a condition for equilibrium uniqueness, although I do not encounter any equilibrium multiplicity in numerical simulations.

With the help of numerical simulations solving for the equilibrium price paths I analyze and disentangle the effects driving the price dynamics. Equilibrium prices can be dispersed, as they depend on the relation of firms' capacities and remaining selling periods, which determine the maximum expected demand. If firms have each enough capacity to serve the entire expected demand, there is harsh competition and firms post identically low but positive prices. However, if one firm cannot serve the entire expected demand by itself, yet the sum of firms' capacities exceeds expected demand, prices are dispersed and under duopoly the larger firm typically allows the smaller firm to undercut because it has a larger reservation value, expecting to sell as a monopolist as soon as the smaller firm is out of stock. Contrary, if the sum of both firms' capacities is lower than expected demand, i.e. if capacities are scarce, the larger firm undercuts. This shows that price-leadership is not monotone in capacity-leadership. Further, I find that under competition prices might be neither monotone in a firm's capacity nor in remaining time, ceteris paribus. However, if capacities are scarce or if there is a monopoly, prices weakly decrease in capacity and weakly increase in remaining time.

Many results in this work are compared to Dudey's (1992) model of dynamic Bertrand pricing.[2] There, all firms can sustain monopoly prices throughout the entire selling time as long as at least one firm cannot serve all expected demand.[3] In my model consumer heterogeneity on the one hand intensifies competition to attract consumers and therefore yields lower and more dispersed prices, unless the market is already very competitive. On the other hand heterogeneity increases firms' market power as goods become less homogeneous, such that equilibrium prices can increase in heterogeneity.

I consider three policy discussions. For these, average prices, expected profits, consumer surplus and total sales as a proxy for efficiency are computed for expected equilibrium price paths. First, I analyze policies allowing consumers to become more forward-looking. Facing an inter-temporal problem, consumers trade off buying at current prices and waiting for possibly better prices, however at the risk of being rationed. Although there is an overall efficiency loss (less sales), consumers are still better off when forward-looking because prices (and hence total industry profits) are depressed. Second, many industries using dynamic pricing techniques are frequently in the center of competition policy discussions, e.g. airline mergers. Results from this model suggest that competition policy is especially relevant if overall market capacities are excessive relative to expected

[2] In this model with unit Bertrand demand per period there is no consumer heterogeneity.

[3] Only if capacities are symmetric, below-marginal cost pricing is possible.

demand. Contrary, under capacity scarcity market power does not have a significant effect on average prices and welfare measures. Third, I analyze the equilibrium behavior of firms in the ex-ante game of capacity production. In equilibrium, capacity production can be excessive relative to expected demand and more capacity than under capacity production collusion and under monopoly is built. Nevertheless, from an efficiency perspective this is still too little capacity, resulting in higher dynamic price paths which lead to inefficiently low sales.

The remainder of this paper is structured as follows: Section 2 relates this paper's contribution to the literature. Subsequently, Section 3 introduces the basic model of myopic consumers as well as the equilibrium analyses, while the model with forward-looking consumers is found in Section 4. With the help of simulations, in Section 5 I present comparative statics results for many model parameters. In Section 6 I provide welfare and policy discussions before Section 7 concludes. Appendix A contains all proofs as well as auxiliary derivations and results in order of appearance in the main text, whereas the equilibrium simulation programming codes are delegated to Appendix B. Note for the whole paper that whenever findings are from simulations I refer to them as *Results*, while I denote them *Lemma* or *Proposition* when proven analytically.

2 Literature

Initiated by Gallego and Van Ryzin (1994), the traditional revenue management literature typically considers a monopolist's problem when selling multiple units of the same product to consumers arriving sequentially and possibly with some uncertainty. Much work considering differentiated or forward-looking consumers has been done, yet primarily in models of monopoly, whereas I consider competition. For example, Su (2007) studies the case of consumers being differentiated with respect to their valuation and their patience. Dilme and Li (2016) study the situation with differentiated consumers, where a seller can periodically use fire-sales in order to reduce capacity and be able to commit to higher future prices for high type consumers. Board and Skrzypacz (2016) characterize deterministic cutoffs, depending on inventory and remaining time, also under monopoly. Hörner and Samuelson (2011) consider the monopoly problem of selling to forward-looking consumers with independent private values, giving rise to a trade-off between imperfect price discrimination and setting a "reserve" price. Then, the monopolist's price continuously decreases in time and only jumps up after a sale. Meisner (2017) builds upon this work but extends to competition. He shows that if capacity is scarce, the competitive price path is identical to the one of the monopolist. I consider homogeneous consumer valuations with heterogeneous taste shocks (logit demand) under competition and allow for other capacity constellations than scarcity. Nevertheless, under capacity scarcity my model predicts similarity between the monopoly and competitive price, too. Other work, e.g. Nocke and Peitz (2007) and Möller and Watanabe (2010), make use of a mechanism design approach, studying a monopolist who discriminates inter-temporally between low and high type consumers, and thereby comparing uniform pricing to advance purchase discounts and clearing sales. For an excellent overview of earlier literature, see Talluri and Van Ryzin (2006). Summing up this paragraph, my work adds to the literature in the sense that I consider oligopolistic competition with forward-looking consumers and general capacity constellations.

My work goes back to the Bertrand-Edgeworth (Edgeworth et al., 1925) model, i.e. Bertrand competition with capacity constrains. With static pricing (open-loop with constant prices) a pure-strategy equilibrium may not exists if for instance in a duopoly both firms together can serve the entire market but one firm alone does not have enough capacity to serve all demand itself. In these situations different mixed-strategy equilibria with continuous price support exist, depending on the rationing assumptions (Allen and Hellwig, 1986; Kreps and Scheinkman, 1983; Levitan and Shubik, 1972). In my model this is not relevant since I consider dynamic pricing.

Dynamic pricing allows firms to adjust prices at each point in time before the deadline (closed-loop pricing). Dudey (1992) shows that with dynamic instead of static pricing for Bertrand competition a unique pure-strategy equilibrium exists. Depending on capacities relative to remaining selling periods either the monopoly or the perfectly competitive price results, unless both firms have identical capacity stocks, then prices can even be below marginal costs. Furthermore, in the range in which mixed-strategy equilibria occur in static models, in the dynamic model the larger firm will allow the firm with less capacity to sell out at the monopoly price, as it prefers to wait until it remains as the monopolist itself. Contrary, in my model consumers are heterogeneous and therefore firms set dispersed prices such that due to consumers' taste heterogeneity either of the firms may sell. However, as a limit case when consumer heterogeneity tends to zero, I obtain Dudey's (1992) results.

Martínez-de Albéniz and Talluri (2011) adopt Dudey's (1992) model and include demand uncertainty. They show that continuation payoffs determine prices and again the seller with lowest capacity sells her entire stock first, at the price of the reservation value of the next smallest seller, thereby giving rise to a path of volatile prices. Further models with demand uncertainty are for instance Deneckere and Peck (2012), who consider perfect competition and forward-looking consumers, or Dana Jr (1999), who explores price dispersion due to demand uncertainty and market structure. In my work firms' reservation values are crucial, too, but I do not consider any further demand uncertainty, only consumer heterogeneity.

In my model consumers are heterogeneous in the sense that they obtain an idiosyncratic taste shock with respect to each firm's product. This gives rise to the (multinomial) logit demand system (McFadden et al., 1973). With this in hand I can construct states, state transitions and the inter-temporal optimization problem of firms, building upon results from the computational industrial organization and operations research literature. Here, I borrow from Pakes and McGuire (1992) and Doraszelski and Satterthwaite (2010) in the construction of the algorithm finding subgame-perfect equilibria. Perakis and Sood (2006) consider a similar approach, however with an open-loop pricing procedure. In a setting similar to mine, however without forward-looking consumers, Lin and Sibdari (2009) compare complete information to a situation where firms do not know the real-time inventory levels of competitors. In their model there is further demand uncertainty, nevertheless the baseline set-up is similar to mine. Sinitsyn (2008) characterizes mixed-strategy equilibria consisting of a finite number of prices, when there is logit demand but static instead of dynamic pricing.

Of course, there is a large empirical literature considering price dispersion and price discrimination in markets such as the one I study. Notably, Borenstein and Rose (1994) and Gerardi and Shapiro (2009) investigate the effects of airline competition on price dispersion. Escobari (2012) studies dynamic pricing of inventories with uncertain demand and his estimates suggest that prices increase as the inventory level decreases, while they decrease if the deadline approaches, which is in line with my findings. An overview of dynamic pricing results and revenue losses from mis-anticipating consumer behavior, especially forward-looking behavior of consumers, can be found in Gönsch et al. (2013). Sweeting (2012) uses evidence from the secondary market of major league baseball tickets to show that prices decline towards the deadline, and that dynamic pricing methods increase revenues substantially. In a recent paper, Williams (2017) examines dynamic airline pricing considering how fares respond to time and to remaining capacity levels. Using a discrete choice model for consumer demand, he finds that the interaction of these two effects is crucial for pricing, which corresponds with the theoretical predictions of my model.

3 Baseline Model

In this section I will introduce the baseline model with myopic consumers and provide equilibrium derivations.

3.1 Model Specifics

Set-up

Time is discrete and there are $T+1$ total time periods, such that goods can be sold during the T periods and are assumed to perish at the deadline $T+1$. A period is defined by the number of remaining time periods, i.e. $t \in \{T, T-1, ..., 1, 0\}$, where $t = 0$ represents the deadline when no more sales are possible. Each firm $i \in \{1, ..., n\}$ holds $x_i \in \mathbb{N}_{\geq 0}$ capacities at a certain time, with $\mathbf{x} = (x_1, ..., x_n)$. Let $\boldsymbol{\omega} = (t, \mathbf{x})$ define the state of the world and be common knowledge among the firms. In each period, a firm i posts a price $p_i(\boldsymbol{\omega}) \in \mathbb{R}$, s.t. $\mathbf{p}(\boldsymbol{\omega}) = (p_1(\boldsymbol{\omega}), ..., p_n(\boldsymbol{\omega}))$. When selling a unit, this firm's capacity is reduced by one. During the game firms can neither build more capacity nor can they destroy capacity.

The timing in a period is such that first a consumer arrives into the market. Then all *active* firms defined by the set $J(\boldsymbol{\omega}) := \{j \in \{1, ..., n\} \,|x_j \geq 1\}$, i.e. all firms with strictly positive remaining capacity, simultaneously post prices.[4] Subsequently, the consumer, who is assumed to be myopic in the baseline model, chooses to buy from one of these firms $j \in J(\boldsymbol{\omega})$ or to exit the market, denoted by the outside option $j = 0$.[5] Consequently, at most one trade per period/state takes place. Depending on this consumer choice, the state transits to the next state in period $t-1$, as defined by the following state transition.

State Transitions

Given the consumer choice $j \in J(\boldsymbol{\omega}) \cup \{0\}$ in state $\boldsymbol{\omega} = (t, \mathbf{x})$, the state transition function into a next state $\boldsymbol{\omega'} = (t', \mathbf{x}')$ reads

$$\boldsymbol{\omega'} |\, (\boldsymbol{\omega}, j) = \begin{cases} (t-1, \mathbf{x}) & \text{if } \; j = 0, \\ (t-1, \mathbf{x} - \mathbf{e}_j) & \text{if } \; j \in J(\boldsymbol{\omega}), \end{cases} \tag{1}$$

[4] Alternatively one could assume that firms without capacity "set" a price of infinity.

[5] Myopic consumer do not need knowledge of the state as they only consider the posted prices. Contrary, forward-looking consumers (see Section 4) need this knowledge as they can choose to wait.

where $\mathbf{e}_j$ is a $1 \times n$ vector with all zeros but the jth entry equal to 1. In words, the state transits to the next period $t-1$ with unchanged capacities $\mathbf{x}$ if the consumer chose not to buy, i.e. $j = 0$, or to the next period $t-1$ after she bought from firm j with $x_j \geq 1$, such that this firm j's capacity is then reduced by one. Note that time cannot decrease further than to its deadline because in $t = 0$ the game ends. Consumers cannot choose to buy from a firm with zero capacity, i.e. $x_j = 0$, as that firm would be inactive.

Demand

In each period one consumer arrives in the market with probability one.[6] Firms offer homogeneous products of quality v that expire at the same time $t = 0$. The consumer receives utility from the good's quality v and price p_i, depending on where she buys. Assuming linear utility in product quality and price, buying from firm $i \in J(\boldsymbol{\omega})$ gives the consumer utility[7]

$$U_i = v - p_i + \mu\, \epsilon_i, \tag{2}$$

while the outside option, i.e. exiting the market without buying, grants the normalized utility of

$$U_0 = 0 + \mu\, \epsilon_0. \tag{3}$$

A consumer receives idiosyncratic preference shocks ϵ_j for each choice $j \in J(\boldsymbol{\omega}) \cup \{0\}$, which are assumed to be drawn independently from the Extreme Value Type-I (or Gumbel) distribution and are not observed by the sellers.[8] These shocks are weighted with $\mu > 0$, the measure of consumer heterogeneity, which is common knowledge of the firms. Following the discrete choice model, a consumer chooses $i \in J(\boldsymbol{\omega}) \cup \{0\}$ if and only if

$$U_i(v, p_i, \mu, \epsilon_i) \geq U_j(v, p_j, \mu, \epsilon_j),\ \forall j \in J(\boldsymbol{\omega}) \cup \{0\} \text{ and } j \neq i.$$

This yields the multinomial logit demand system (see e.g. Train (2009) for the derivation of the discrete choice probabilities), such that the demand probability that a consumer chooses product $i \in J(\boldsymbol{\omega})$ over all other choices is given by

[6] In Section 5.5 I consider general arrival rates.

[7] In the standard discrete choice model v and p_i can be weighted with β and α respectively, i.e. taste parameters over quality and price, however I normalize $\alpha = \beta = 1$. Note that this is without loss of generality because by adjusting v and μ we can essentially change α and β.

[8] The distribution function is $f(\epsilon_j) = \exp(-\epsilon_j)\, \exp\left[-\exp(-\epsilon_j)\right]$.

$$D_i\Big(\mathbf{p}(\boldsymbol{\omega}),\boldsymbol{\omega}\Big) = \frac{\exp\left(\frac{v-p_i(\boldsymbol{\omega})}{\mu}\right)}{1+\sum\limits_{J(\boldsymbol{\omega})}\exp\left(\frac{v-p_j(\boldsymbol{\omega})}{\mu}\right)}, \tag{4}$$

while inactive firms' demand is zero, i.e. $D_i(\mathbf{p}(\boldsymbol{\omega}),\boldsymbol{\omega}) = 0$ if $x_i = 0$. Further, the consumer chooses the outside option with probability

$$D_0\Big(\mathbf{p}(\boldsymbol{\omega}),\boldsymbol{\omega}\Big) = \frac{1}{1+\sum\limits_{J(\boldsymbol{\omega})}\exp\left(\frac{v-p_j(\boldsymbol{\omega})}{\mu}\right)}. \tag{5}$$

Note that the sum over all demand probabilities including the outside option adds up to one, i.e. $\sum_j D_j(\mathbf{p}(\boldsymbol{\omega}),\boldsymbol{\omega}) = 1$, with $j \in J(\boldsymbol{\omega}) \cup \{0\}$. Further note that logit demand, in contrast to simple Bertrand unit demand, does not entail demand curve jumps, resulting in non-trivial intermediate prices rather than only extreme monopoly or marginal cost pricing. Selling probabilities are also more dispersed, in the sense that any firm might sell in a given period with a certain probability, giving rise to the volatile price dynamics.

For $\mu \to 0$ consumers become homogeneous and we tend to the (unit) Bertrand world, where $U_i = v - p_i$. Then demand also tends to its Bertrand equivalent. See Appendix A.1 for these derivations and Section 5 for comparative statics in consumer heterogeneity μ. Further note that $D_0(\mathbf{p})$ increases in μ if $p_j < v$ for all $j \in J(\boldsymbol{\omega})$, while $D_i(\mathbf{p})$ decreases in μ if $p_i \leq p_j$ and $p_i < v$ for all $j \neq i$. Intuitively, here heterogeneity increases the outside option utility relatively to the utility from a good.

Firms' Problem

In each period (state), each active firm $i \in J(\boldsymbol{\omega})$ chooses a price $p_i(\boldsymbol{\omega})$ for its product to maximize $V_i(\mathbf{p}(\boldsymbol{\omega}),\boldsymbol{\omega})$, the expected net present value of current and all future payoffs, depending on all firms' price choices $\mathbf{p}(\boldsymbol{\omega})$ in this state. Note that marginal costs are normalized to zero. If a firm i's capacity is zero, it is inactive and its value function is equal to zero. In the deadline firms' value function is zero, too, such that $V_i(\mathbf{p}(\boldsymbol{\omega}),\boldsymbol{\omega}) = 0$ if $x_i = 0$ or $t = 0$. If $x_i \geq 1$ and $t \geq 1$, then a firm's value function is given by

$$V_i(\mathbf{p}(\boldsymbol{\omega}),\boldsymbol{\omega}) \quad = \quad D_i\Big(\mathbf{p}(\boldsymbol{\omega}),\boldsymbol{\omega}\Big)p_i(\boldsymbol{\omega}) + \delta\sum_j\Big[D_j\Big(\mathbf{p}(\boldsymbol{\omega}),\boldsymbol{\omega}\Big)W_{i,j}(\boldsymbol{\omega})\Big], \tag{6}$$

where δ is the discount factor[9] and

$$W_{i,j}(\boldsymbol{\omega}) = V_i\Bigg(\mathbf{p}^*\Big(\boldsymbol{\omega}'|\,(\boldsymbol{\omega},j)\Big),\boldsymbol{\omega}'|\,(\boldsymbol{\omega},j)\Bigg) \tag{7}$$

is the expectation of the value function of firm i for the next period, where all firms will post equilibrium prices, conditional on the consumer having chosen $j \in J(\boldsymbol{\omega})\cup\{0\}$ now.[10] Hence a firm's value function (equation 6) consists of current-period payoffs, i.e. demand $D_i(\mathbf{p}(\boldsymbol{\omega}),\boldsymbol{\omega})$ multiplied with price $p_i(\boldsymbol{\omega})$, as well as the sum over next period's expected value functions $W_{i,j}(\boldsymbol{\omega})$, weighted with the probabilities $D_j(\mathbf{p}(\boldsymbol{\omega}),\boldsymbol{\omega})$ that these states are reached. The game will be solved by backward induction. As a boundary condition, in the last period $t = 1$ a firm's value function consists only of current-period payoffs, i.e. $V_i(\mathbf{p}(\boldsymbol{\omega}),\boldsymbol{\omega}) = D_i(\mathbf{p}(\boldsymbol{\omega}),\boldsymbol{\omega})p_i(\boldsymbol{\omega})$ if $t = 1$.

The best-response strategy of firm i in a state $\boldsymbol{\omega}$, given competitor prices $\mathbf{p_{-i}}(\boldsymbol{\omega})$, is

$$p_i\Big(\mathbf{p_{-i}}(\boldsymbol{\omega}),\boldsymbol{\omega}\Big) \;=\; \arg\max_{p_i} V_i\Big(\mathbf{p}(\boldsymbol{\omega}),\boldsymbol{\omega}\Big). \tag{8}$$

Differentiating the value function in (6) w.r.t. $p_i(\boldsymbol{\omega})$ and considering the demand derivatives as in Appendix A.1,[11] gives

$$\frac{\partial V_i(\mathbf{p},\boldsymbol{\omega})}{\partial p_i} \;=\; D_i\left(\mathbf{p},\boldsymbol{\omega}\right)\Big[1-\frac{1}{\mu}p_i-\frac{\delta}{\mu}W_{i,i}(\boldsymbol{\omega})+\frac{1}{\mu}V_i(\mathbf{p},\boldsymbol{\omega})\Big] \tag{9}$$

$$\;=\; D_i\left(\mathbf{p},\boldsymbol{\omega}\right)\left[1-\frac{1}{\mu}\sum_{j\neq i}D_j\left(\mathbf{p},\boldsymbol{\omega}\right)\Big[p_i+\delta\,W_{i,i}(\boldsymbol{\omega})-\delta\,W_{i,j}(\boldsymbol{\omega})\Big]\right]. \tag{10}$$

Further differentiating gives

$$\frac{\partial^2 V_i(\mathbf{p},\boldsymbol{\omega})}{\partial p_i^2} \;=\; \frac{1}{\mu}\frac{\partial V_i(\mathbf{p},\boldsymbol{\omega})}{\partial p_i}\Big[2D_i\left(\mathbf{p},\boldsymbol{\omega}\right)-1\Big]-\frac{1}{\mu}D_i\left(\mathbf{p},\boldsymbol{\omega}\right). \tag{11}$$

[9]The firms' discount factor could be normalized to $\delta = 1$ as selling periods in these markets typically last only a few months.

[10]The continuation values only depends on the current state $\boldsymbol{\omega}$, i.e. are independent of the firms' actions (prices) in the current period. Current prices only determine the probabilities with which these continuations are reached. Future equilibrium prices are anticipated and hence taken as given in a period. Note that I show equilibrium existence by backward induction (and discuss uniqueness) in the next subsection.

[11]Note that in the following I drop some arguments $(\boldsymbol{\omega})$ for simplicity.

3.2 Equilibrium Existence

This is a multi-period game, therefore the equilibrium concept will be Sub-game Perfect Nash. For the existence proof we need the following two lemmas.

Lemma 1. *The best-response function $p_i(\mathbf{p_{-i}}(\boldsymbol{\omega}), \boldsymbol{\omega})$ is bounded, i.e.* $\inf p_i(\mathbf{p_{-i}}(\boldsymbol{\omega}), \boldsymbol{\omega}) > -\infty$ *and* $\sup p_i(\mathbf{p_{-i}}(\boldsymbol{\omega}), \boldsymbol{\omega}) < \infty$.

Intuitively, the best-response price $p_i(\mathbf{p_{-i}}(\boldsymbol{\omega}), \boldsymbol{\omega})$ in equation (8) is bounded because very low and very high prices are never a best response. For this I show that the change in the expected value function $V_i(\mathbf{p}(\boldsymbol{\omega}), \boldsymbol{\omega})$ becomes positive for low prices, i.e. a condition for $\partial V_i(\mathbf{p}(\boldsymbol{\omega}), \boldsymbol{\omega})/\partial p_i > 0$, and that the change in $V_i(\mathbf{p}(\boldsymbol{\omega}), \boldsymbol{\omega})$ becomes negative for high prices, i.e. a condition for $\partial V_i(\mathbf{p}(\boldsymbol{\omega}), \boldsymbol{\omega})/\partial p_i < 0$. This lemma is necessary for the construction of a compact strategy set in the existence proof.

Lemma 2. *The expected value function $V_i(\mathbf{p}(\boldsymbol{\omega}), \boldsymbol{\omega})$ is quasi-concave in p_i.*

The second lemma is necessary for the existence of a unique best-response of firm i for each $\mathbf{p_{-i}}$. With these two lemmas we can state the main proposition of equilibrium existence.

Proposition 1. *There exists a pure-strategy Sub-game Perfect Nash Equilibrium.*

In the proof for this proposition I follow Fudenberg and Tirole (1991) to show that in each sub-game an equilibrium exists. Then, by backward induction there exists a pure-strategy Sub-game Perfect Nash Equilibrium for the whole game. For this note that in the final period $t = 1$ no continuation values need to be considered.

To characterize the equilibrium in a given state $\boldsymbol{\omega}$, from equation (10) the first-order condition of firm i implicitly defining an equilibrium price choice $p_i^*(\boldsymbol{\omega})$ is

$$\mu = \sum_{j \neq i} D_j\Big(\mathbf{p}^*(\boldsymbol{\omega}), \boldsymbol{\omega}\Big)\Big[p_i^*(\boldsymbol{\omega}) + \delta W_{i,i}(\boldsymbol{\omega}) - \delta W_{i,j}(\boldsymbol{\omega})\Big], \tag{12}$$

where $j \in J(\boldsymbol{\omega}) \cup \{0\}$. Intuitively, a firm will post a price reflecting the trade-off between current-period profit $p_i^*(\boldsymbol{\omega})$ together with the expected continuation value upon selling itself, $\delta W_{i,i}(\boldsymbol{\omega})$, and its reservation value which is the continuation value after the consumer chooses the other j, i.e. $\delta W_{i,j}(\boldsymbol{\omega})$. A similar point is made by Martínez-de Albéniz and Talluri (2011) for dynamic Bertrand competition.

The equilibrium valuation function in a given state $\boldsymbol{\omega}$ then becomes

$$V_i(\mathbf{p}^*(\boldsymbol{\omega}), \boldsymbol{\omega}) \quad = \quad p_i^*(\boldsymbol{\omega}) + \delta W_{i,i}(\boldsymbol{\omega}) - \mu. \tag{13}$$

3.3 Equilibrium Uniqueness

In this section I will state a sufficient condition for uniqueness of the sub-game perfect equilibrium under duopoly. Moreover, I will show analytically that this condition holds for important parameter constellations, while for general parameter levels I use a simulation approach.

For the following condition I restrict attention to duopoly situations. A monopolist's price choice is unique, as already established in the last paragraph of the proof of Proposition 1.

Condition 1. *There exists a unique pure-strategy Sub-game Perfect Nash Equilibrium under duopoly, if for all* $i, j \in \{1, 2\}$ *with* $j \neq i$

$$\left.\frac{\partial V_i(\mathbf{p}, \boldsymbol{\omega})}{\partial p_j}\right|_{\frac{\partial V_i(\mathbf{p},\boldsymbol{\omega})}{\partial p_i}=0} \in (-1,\, 1). \tag{14}$$

While this condition is rather technical and without further intuition, I can show that for $\mu \to 0$ Condition 1 holds and the equilibrium is unique, which is in line with the uniqueness result in Dudey (1992).

Proposition 2. *There exists a unique pure-strategy Sub-game Perfect Nash Equilibrium under duopoly for* $\mu \to 0$*.*

Moreover, for $\mu > 0$ I can show that the Nash equilibrium in the last period $t = 1$ is unique. Here, no more continuation values are relevant. Similarly, whenever $\min_i x_i \geq t$ and $i \in J(\boldsymbol{\omega})$, i.e. when the active firms have individual excess capacities (compare Section 5.1), all continuation values of all firms are equal. Independent of the choice of the consumer all firms will have sufficient capacity until the deadline, such that uniqueness can be shown analogously.

Proposition 3. *Under duopoly, there exists a unique pure-strategy Sub-game Perfect Nash Equilibrium if* $\min_i x_i \geq t$ *and* $i \in J(\boldsymbol{\omega})$*, and all firms post the same price.*

To see that for all t where $\min_i x_i \geq t$ all active firms will post identical prices, note that here the first-order condition from (12) simplifies to

$$\mu = \Big[1 - D_i\Big(\mathbf{p}^*(\boldsymbol{\omega}), \boldsymbol{\omega}\Big)\Big]\; p_i^*(\boldsymbol{\omega}), \tag{15}$$

for all active firms i, which is (implicitly) solved by a unique and symmetric price for all i because the demand function at given prices is identical in all these t and all firms are symmetric in the sense that they each hold capacity $x_i \geq t$.[12] Intuitively, firms treat all periods with individual excess capacities like the final period and post the *final-period* price. Under competition, I will refer to this price as the *total-competition* price, since it defines the lower-bound for competitive prices.[13] Naturally, under monopoly there exists such a lower-bound final-period price, too, which is posted whenever $x_i \geq t$.

Finally, for higher levels of $\mu > 0$ in all t where $\min_i x_i < t$ I provide the following result, which is based on simulations, where I do not encounter equilibrium multiplicity for any considered parameter constellation while looping over a wide range of starting values for the root-finding algorithm. The program for this is provided in Appendix B.

Result 1. *There exists a unique pure-strategy Sub-game Perfect Nash Equilibrium under duopoly.*

4 Forward-looking Consumers

In this section I will consider forward-looking consumers and present the model specifics as well as equilibrium results. Effects of a policy allowing consumers to be forward-looking will be discussed in Section 6.2.

Contrary to myopic consumers, forward-looking consumers have knowledge of the state of the world, which next to the number of remaining time periods and firms' capacity levels will now also consist of the number of consumers in the market.[14] Furthermore, I assume that in each time period a consumer can choose between buying from any of the active firms or waiting in the market for at least one more period, rather than exiting the market. Nevertheless, waiting in every period until the deadline provides the same (outside option) utility for forward-looking consumers as exiting for myopic consumers.

4.1 Model

Set-up

As in the baseline model, in each period one new consumer arrives. Since consumers can now choose to wait, more than one consumer might "accumulate" in the market, such

[12] Note that if μ is not too (insensibly) high, this price is lower than consumer valuation v. Refer to the discussion of the comparative statics of μ in Section 5.3 for more details.

[13] Only under symmetric aggregate capacities even lower prices are possible (see Section 5.1).

[14] One could also consider a third, "intermediate" case, where consumers are not informed about capacities, but know the remaining time, allowing them to be still fully forward-looking. This case would require more complex analysis involving consumers having beliefs about capacities and possibly firms using prices as capacity signals.

that at the deadline at most as many consumers as total periods can be in the market. Let $c \in \mathbb{N}_+$ be the number of consumers in the market at a certain time. Again, I assume that in each period only one trade can occur, i.e. at most one of possibly several waiting consumers can buy from only one firm. The remaining consumers are "rationed" and have to wait for (at least) one more period.

Then the timing in a period is such that first a new consumer arrives into the market with probability one, while all consumers who did not buy in the previous period are in the market as well. Second, firms simultaneously post prices and a lottery draws out of all consumers in the market one *lucky* consumer.[15] The probability to be the lucky consumer is equal to $1/c$, while the probability to be rationed is equal to $(c-1)/c$. Third, the lucky consumer chooses to buy from an active firm or to further wait in the market. Possible trade between the lucky consumer and a firm takes place and the period ends. All consumers who did not buy will remain in the market for the next period.

State Transition

The state is now given by the vector $\boldsymbol{\omega} = (t, \mathbf{x}, c)$, i.e. it additionally features the number of consumers c in the market, and is common knowledge among all firms and consumers.

Depending on the choice of the lucky consumer, the state transition function from $\boldsymbol{\omega} = (t, \mathbf{x}, c)$ into a next state $\boldsymbol{\omega'} = (t', \mathbf{x}', c')$ is

$$\boldsymbol{\omega'} \,|\, (\boldsymbol{\omega}, j) = \begin{cases} (t-1, \mathbf{x}, c+1) & \text{if } j = 0, \\ (t-1, \mathbf{x} - \mathbf{e}_j, c) & \text{if } j \in J(\boldsymbol{\omega}), \end{cases} \tag{16}$$

where $\mathbf{e}_j$ is a $1 \times n$ vector with all zeros but the jth entry equal to 1. If the consumer waits ($j = 0$), $\mathbf{x}$ capacities remain and one additional new consumer arrives, such that there will be $c + 1$ consumers in the next period $t - 1$. If the consumer chooses firm $j \in J(\boldsymbol{\omega})$ with $x_j \geq 1$, then this firm j's capacity is reduced by 1, while (possibly) unserved consumers as well as the one newly arriving consumer, i.e. $c - 1 + 1 = c$, will be in the market in the next period. Note that time cannot decrease past its deadline boundary $t = 0$.

Consumer Utility and Demand

In the world of forward-looking consumers waiting gives the consumer utility of $\delta_c\, W_c(\boldsymbol{\omega})$, where $W_c(\boldsymbol{\omega})$ is the consumer's expected valuation of waiting which is introduced in the

[15] Note that the order of the consumer lottery and firms' price setting is irrelevant.

next paragraph. Let $\delta_c \in [0, \delta]$ be the discount rate for all consumers, which could be smaller than the firms' discount rate δ.[16] I will interpret δ_c as consumers' degree of patience, i.e. the degree to which they are forward-looking (see Section 5.4 for a discussion). Note that for $\delta_c = 0$ the model becomes identical to the baseline model of myopic consumers. The lucky consumer chooses from $J(\boldsymbol{\omega}) \cup \{0\}$, where choices $j \in J(\boldsymbol{\omega})$ exhibit the products of the active firms, while $j = 0$ represents the choice of waiting. Utility of the lucky consumer when choosing an active firm's product i is again $U_i = v - p_i + \mu\,\epsilon_i$. Waiting gives the consumer utility of[17]

$$U_0 = \mu\,\delta_c\,W_c(\boldsymbol{\omega}) + \mu\,\epsilon_0. \tag{17}$$

All choice options are again associated with independent Extreme Value Type-I distributed taste shocks ϵ_j. Then the multinomial logit demand system yields the probabilities that the consumer chooses a product $i \in J(\boldsymbol{\omega})$ or waits (0),

$$D_i(\mathbf{p}(\boldsymbol{\omega}), \boldsymbol{\omega}) = \frac{\exp\left(\frac{v - p_i(\boldsymbol{\omega})}{\mu}\right)}{\exp\left[\delta_c\,W_c(\boldsymbol{\omega})\right] + \sum\limits_{J(\boldsymbol{\omega})} \exp\left(\frac{v - p_j(\boldsymbol{\omega})}{\mu}\right)}, \tag{18}$$

$$D_0(\mathbf{p}(\boldsymbol{\omega}), \boldsymbol{\omega}) = \frac{\exp\left[\delta_c\,W_c(\boldsymbol{\omega})\right]}{\exp\left[\delta_c\,W_c(\boldsymbol{\omega})\right] + \sum\limits_{J(\boldsymbol{\omega})} \exp\left(\frac{v - p_j(\boldsymbol{\omega})}{\mu}\right)}, \tag{19}$$

while inactive firms obtain zero demand, i.e. $D_j(\mathbf{p}(\boldsymbol{\omega}), \boldsymbol{\omega}) = 0$ if $x_j = 0$.

Consumers' Problem

In each period the lucky consumer chooses between buying from a firm or waiting for (at least) one more period in the market. Thereby she considers the following trade-off. On the one hand, she could purchase at current prices from one of the active firms with certainty. On the other hand, she could wait, expecting that firms become less confident about selling their goods and hence will further decrease prices. However, waiting involves the risk of being rationed in the next period(s), amplified by the arrival of new consumers. When the consumer chooses where to buy or whether to wait, she anticipates all possible future prices of all sub-games, which determine her expected valuation of waiting.

[16] Note that δ can be normalized to one.

[17] Note that I multiply the observable part of the outside option utility with μ, too, such that the ratio to the shock ϵ_0 remains constant, as in the baseline model, where the observable part of the outside option is zero. This is merely a normalization and hence does not affect the results qualitatively.

Given the current state $\boldsymbol{\omega}$ and considering future rationing risks, let the transition probability matrix of the lucky consumer be $\Phi(\boldsymbol{\omega})$, a matrix of three dimensions, $T \times \boldsymbol{X} \times C$, with entries in $[0, 1]$.[18] This matrix shall state for each entry, i.e. for each possible state $\boldsymbol{\omega}' = (t', \mathbf{x}', c') \in \{\{0, .., T\} \times \{0, .., \boldsymbol{X}\} \times \{1, .., C\}\}$, the probability that in that state $\boldsymbol{\omega}'$ the currently lucky consumer will be for the first time again the lucky consumer, given that she decides to wait in the current state $\boldsymbol{\omega}$.[19] In Appendix A.2 I provide the details of this transition matrix $\Phi(\boldsymbol{\omega})$.[20] In order to solve for the equilibrium it is important to note that the consumer transition matrix $\Phi(\boldsymbol{\omega})$ only depends on the current state $\boldsymbol{\omega}$. All future sub-games' prices, demand decisions and rationing probabilities are fully anticipated in expectation by consumers as they only depend on the current state.

Next, let $\Omega(\boldsymbol{\omega})$ be a matrix of three dimensions, $T \times \boldsymbol{X} \times C$, where the entries are equal to the lucky consumer's value function in all possible states, as expected after choosing to wait in $\boldsymbol{\omega}$. All entries of $\Omega(\boldsymbol{\omega})$ at the deadline, i.e. for $t' = 0$, shall be equal to zero as all goods will have perished at the deadline. Each other entry $(t', \mathbf{x}', c') \in \{\{0, .., T\} \times \{0, .., \boldsymbol{X}\} \times \{1, .., C\}\}$ is equal to the consumer's value function in that state,

$$\Omega(\boldsymbol{\omega})_{[t', \boldsymbol{x}', c']} \quad = \quad V_c(t', \boldsymbol{x}', c'). \tag{20}$$

The value function of the lucky consumer in a state $\boldsymbol{\omega}$ is equal to her consumer surplus as derived for the logit model e.g. in the textbook by Train (2009) and reads[21]

$$V_c(\boldsymbol{\omega}) \quad = \quad \ln \left\{ \exp\left[\delta_c \, W_c(\boldsymbol{\omega})\right] + \sum_{J(\boldsymbol{\omega})} \exp\left(\frac{v - p_j(\boldsymbol{\omega})}{\mu} \right) \right\}. \tag{21}$$

The expected valuation of a consumer upon selecting to wait shall be then given by

$$W_c(\boldsymbol{\omega}) \quad = \quad \sum_{t'} \sum_{\boldsymbol{x}'} \sum_{c'} \left[\Phi(\boldsymbol{\omega})_{[t', \boldsymbol{x}', c']} \cdot \Omega(\boldsymbol{\omega})_{[t', \boldsymbol{x}', c']} \right], \tag{22}$$

[18] $C = c + t$ is the maximum number of consumers that can accumulate until the end.

[19] For example, consider a state ($t = 2$, $x_1 = 1$, $c = 1$). If the only (and hence lucky) consumer waits, her transition probability matrix consists of zeros for all possible states, with one exception: in state ($t = 1$, $x_1 = 1$, $c = 2$) she will be for the first time again the lucky consumer with probability 1/2, which is the probability that she and not the newly arrived consumer will be drawn to be lucky. If the new consumer is drawn, which happens with probability 1/2, the older consumer does not get to be the lucky consumer again and reaches the deadline, obtaining the outside option of no purchase.

[20] Note that unlucky consumers do not become active until they are drawn by the lottery, hence their transition function is irrelevant to the solution of the problem until then.

[21] Note that this term is the log of the denominator of a choice probability and therefore is often called 'the log-sum term', although this equivalence "has no economic meaning" (Train, 2009).

i.e. the sum of all elements obtained from element-wise multiplying the transition probability matrix with the matrix of value functions. Note that in the last selling period $t = 1$ there are no continuation values, hence waiting grants the outside option valuation $W_c(\boldsymbol{\omega}) = 0$ whenever $t = 1$, completing the recursion. Intuitively, $W_c(\boldsymbol{\omega})$ gives the expected equilibrium-path- and rationing-risk-probability-weighted valuation of waiting.

Firms' Problem

In each state $\boldsymbol{\omega} = (t, \mathbf{x}, c)$, all active firms i simultaneously post their prices $p_i(\boldsymbol{\omega})$ to maximize their value function $V_i(\mathbf{p}(\boldsymbol{\omega}), \boldsymbol{\omega})$. The value function is basically identical to the one in the baseline model as in equation (6). The only difference is that consumers here have the choice of waiting rather than exiting the market. Since in each period only one consumer is selected to choose, the firms' problem remains equal: firms compete against each other for one trade per period, thereby considering all possible continuation values. However, consumers' demand as in equations (18) and (19) now also depends on their valuation of waiting (equation 22), which firms' need to take into account.

4.2 Equilibrium

Equilibrium Existence

The proof of equilibrium existence follows the same line as in the baseline model, hence is only sketched here.

Proposition 4. *There exists a pure-strategy sub-game perfect Nash equilibrium, when consumers are forward-looking.*

In each period firms' price choices are a fix point, as in the case of myopic consumers. The only difference to the proof of Proposition 1 is that continuation values are given for consumer choices including waiting instead of exiting the market. Still, a firm's value function is quasi-concave in its own price and the best-response functions are bounded. In each period only one consumer is drawn to choose between buying or waiting, whereas all other consumers remain inactive for this period. This lucky consumer chooses after firms have posted their prices, whereby she takes current firm prices as given and anticipates all prices and continuation values as well as her transition probabilities for all future sub-games (states). Hence, no consumer choice fix-point argument is necessary but the unique optimal consumer choice directly follows from utility maximization, resulting in the above specified demand system, which is anticipated by firms. Therefore, in each period a Nash equilibrium exists and since the game is finite, there exists a pure-strategy Sub-game Perfect Nash Equilibrium for the whole game.

Equilibrium Uniqueness

Similar to the case of myopic consumers, we can show equilibrium uniqueness under duopoly for $\mu \to 0$ analytically and also for all t where $\min_i x_i \geq t$ and $i \in J(\boldsymbol{\omega})$. The proofs are analogous to the ones in the previous section for Propositions 2 and 3.

Proposition 5. *There exists a unique pure-strategy Sub-game Perfect Nash Equilibrium under duopoly, when consumers are forward-looking and* $\mu \to 0$. *If* $\min_i x_i \geq t$ *and* $i \in J(\boldsymbol{\omega})$, *the duopoly equilibrium is unique for* $\mu > 0$, *too, and all firms post equal prices in a given* t.

The equilibrium price levels are implicitly defined by first-order conditions analogous to (15). However, note that unlike in the myopic consumer case, here firms' prices in all t where $\min_i x_i \geq t$, $i \in J(\boldsymbol{\omega})$, are not equal to the final-period price. While firms here still post unique and symmetric prices in a given period of individual excess capacities, this price is not constant in t. This is because demand depends on consumers' valuation for waiting which increases in t as consumers' rationing risk rises the closer we get to the deadline (compare Section 6.2).

For all other cases with $\mu > 0$ in all t where $\min_i x_i < t$ I verify equilibrium uniqueness in the simulations by looping over a wide range of different starting values for the root-finding algorithm (compare the program in Appendix B).

Result 2. *There exists a unique pure-strategy Sub-game Perfect Nash Equilibrium under duopoly, when consumers are forward-looking.*

5 Comparative Statics and Discussion

In this section I discuss comparative statics to shed light on equilibrium pricing. I study the effects of differences in capacity levels relative to remaining selling periods, in consumer heterogeneity, in consumer patience, in the consumer arrival rate and in the number of firms. Thereby I will distinguish between mainly four different capacity cases (see Section 5.1).

To solve for the equilibrium price paths I employ numerical simulations. In Appendix B I provide the program computing the sub-game perfect equilibrium with forward-looking consumers: Starting in the final period $t = 1$ and then using backward induction, I solve for the equilibrium in each possible state, whereby, given the transition probabilities, all possible sub-games' continuation values are taken into account. Since in a given state the equilibrium prices cannot be solved for in closed-form, I rely on root-finding algorithms. To verify equilibrium uniqueness, I loop over various starting values and do not encounter any multiplicity. Note that whenever I compute representative price paths, trade realizations are a random draw considering the equilibrium demand probabilities. To obtain average price paths I average over 1000 repeated simulations.[22] The findings in this section are based on these equilibrium simulations, whereby I vary parametrization to explore comparative statics and dynamics of the model.

5.1 Capacity Cases

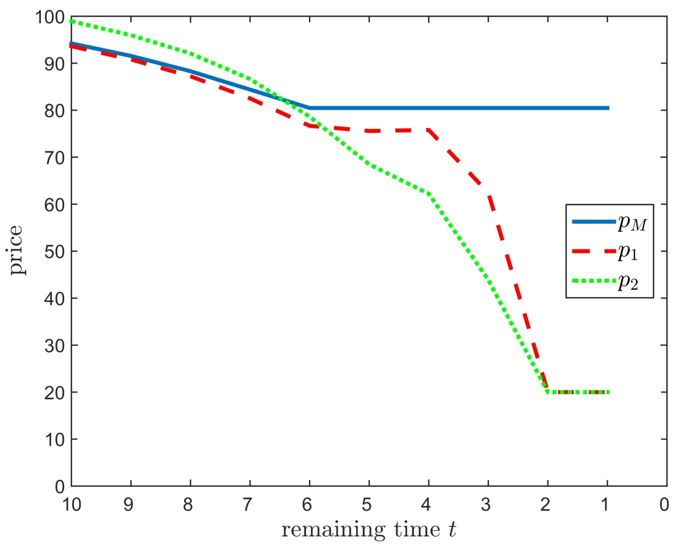
In the pure-strategy sub-game perfect equilibrium different price paths can arise, depending on the relation between firms' capacities and remaining time periods. Note that remaining time periods exhibit an upper bound on expected demand since one consumer arrives per period. In this subsection I will introduce four such capacity cases and show their effect on pricing.

As an initial example consider Figure 2, which shows equilibrium prices with myopic consumers depending on the number of remaining time periods, given fixed capacities of $x_1 = 4$ and $x_2 = 2$ for the entire time under duopoly as well as $x_M = 6$ under monopoly

[22] For all simulations I add to the observable part of utility from a firm i's product $\mu \ln(1/|J(\boldsymbol{\omega})|)$, i.e. μ times the logarithm of the inverse number of active firms. Then, e.g. demand for i becomes $[(1/|J(\boldsymbol{\omega})|)\exp((v - p_i(\boldsymbol{\omega})/\mu)] \,/\, [1 + (1/|J(\boldsymbol{\omega})|)\sum_{J(\boldsymbol{\omega})}\exp((v - p_j(\boldsymbol{\omega})/\mu)]$. Consequently, for forward-looking consumers also the consumer value function $V_c(\boldsymbol{\omega})$ changes accordingly. Like this two duopolists obtain the same aggregate demand as a monopolist if they all post equal prices. Without this modification some price differences between monopoly and competition could be attributed to the fact that consumers in the multinomial-logit model inherently prefer to have more choices, which would have favored competition over monopoly, even at equal prices. Nevertheless, this normalization has no implications for the equilibrium analysis and qualitative comparative statics results of the model parameters.

as a benchmark.[23] In this example we get that $p_1 < p_2$ while $x_1 + x_2 \leq t$, and $p_1 \geq p_2$ while $x_1 + x_2 > t$, whereas $p_1 = p_2$ while $\min\{x_1, x_2\} \geq t$. The monopoly price decreases and remains constant as soon as $t < x_M$.

FIGURE 2: Prices with Fixed Capacities



Parameters: $T = 10$, $v = 100$, $\delta = 1$, $\mu = 10$, myopic consumers. Capacities are fixed for the entire time horizon, s.t. under duopoly $x_1 = 4$, $x_2 = 2$ and under monopoly $x_M = 6$ in all t.

This example already proves that it can be the smaller or the larger firm which posts the lower price. The following proposition summarizes this observation, while more detail is discussed in the subsequent case analysis.

Proposition 6. *Price-leadership is not monotone in capacity-leadership.*

Note that with forward-looking consumers this proposition is robust, as seen in the (respective) example in Figure A.1 of Appendix A.3. For a discussion of why prices with forward-looking consumers rise towards the deadline if all firms can each serve all remaining demand refer to Section 6.2.

Individual Excess Capacities ($\min_i\{x_i\} \geq t,\ i \in J(\boldsymbol{\omega})$)

If all active firms have each enough capacity to serve the entire current and future demand, i.e. firms have individually excess capacities, under duopoly we are in the region

[23] Note for this and all subsequent figures that my model is of discrete capacity levels and time periods. Nevertheless I connect all points in the figures for a better overview.

of $t \leq 2 = x_2$ in Figure 2. Here firms compete harshly and post identically low but positive total-competition prices, equal to the price they would post in the final period. From Proposition 3 we know that under individual excess capacities all firms post the same total-competition price, independent of their capacity. Then, all firms obtain equal demand probabilities such that capacity-leadership can vary throughout the selling horizon. Similarly, a monopolist with excess capacity posts the final-period monopoly price for all remaining time periods, i.e. for $t \leq 6 = x_M$ in Figure 2. Note that in a model of (unit demand) Bertrand competition and static pricing or even dynamic pricing such as in Dudey (1992) firms would price at zero (mark-up) under competition and at the good's value v under monopoly. In my model, due to consumer heterogeneity firms expect dispersed taste realizations. Therefore, larger consumer heterogeneity increases this total-competition price under duopoly.[24]

Asymmetric Aggregate Excess Capacities ($\sum_i x_i > t \wedge \min_i\{x_i\} < t,\ x_i \neq x_j$)

If the sum of firms' capacities exceeds the number of remaining selling periods, but at least one firm cannot serve the entire demand itself, we are in the region $x_2 = 2 < t < 6 = x_1 + x_2$ of Figure 2. This is the capacity case where in the (static) Bertrand-Edgeworth problem mixed-strategy pricing equilibria result. A dynamic model with homogeneous consumers, allowing firms to sell sequentially, such as Dudey (1992), avoids these mixed-strategy equilibria, however the price path is one of constant monopoly prices. In my model with consumer heterogeneity we have a pure-strategy path of dispersed prices, i.e. the smaller or the larger firm could be posting the lower price and either firm could sell. This yields more realistic yet non-trivial intermediate price realizations under competition.[25]

Result 3. *Under asymmetric aggregate excess capacities firms post dispersed prices.*

Typically, the firm with less capacity posts the lower price such that it obtains a higher probability to be chosen by the consumer in this period.[26] In spirit, this result is similar to the one by Dudey (1992), where the lower-capacity firm always undercuts. However, there the result is extreme in the sense that the low-capacity firm can sell out at the monopoly price. The high-capacity firm has an incentive to wait until the smaller firm is out of capacity in order to remain as the monopolist subsequently.

[24] Consider Section 5.3 for details on the effects of consumer heterogeneity.

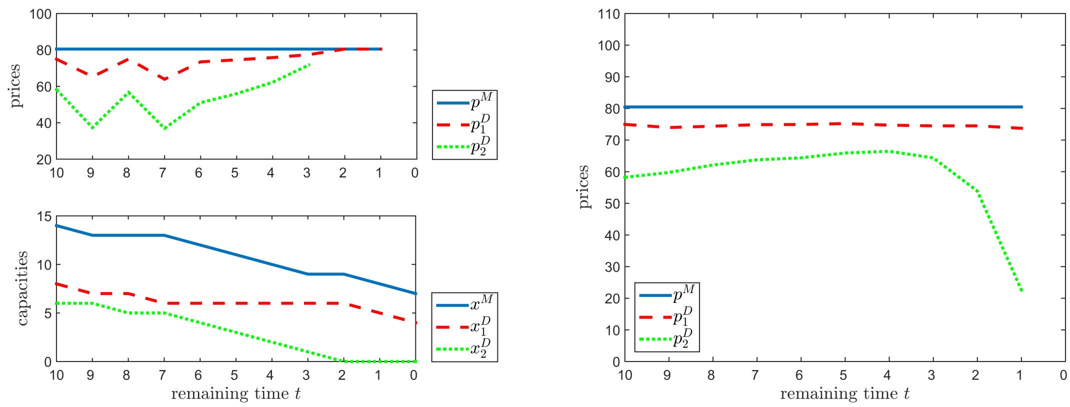
[25] Note that, contrary to Dudey's (1992) dynamic Bertrand model, the static Bertrand model's mixed-strategy prices also yield average prices between the monopoly price and marginal costs.

[26] See Section 5.3 to note that for high values of μ this effect turns around.

Similarly, in my model the smaller firm has the lower reservation value and hence the higher opportunity cost of not selling, which can be best illustrated by the following short example. Consider a situation with two remaining time periods, in which each one consumer arrives. Firm 1 has only one more capacity in stock while firm 2 has two. In the penultimate period firm 1 fears that if no sale occurs or if 2 sells, there will be harsh competition in the last period, i.e. 1 has a low reservation value and consequently undercuts to increase its selling probability. If firm 1 sells now, this would allow firm 2 to sell as a monopolist in the ultimate period instead of engaging in price competition, i.e. firm 2 has a higher reservation value and sets a higher price allowing it to sell only in the case of a favorable consumer taste realization. Heterogeneity of consumers pushes both prices below the monopoly price such that lower taste realizations can be served. This is also true for the larger firm, since by pushing both prices downward the probability of a sale increases and hence it gets more likely that the larger firm becomes monopolist.

FIGURE 3: Asymmetric Aggregate Excess Capacities

(a) Representative Price and Capacity Path (b) Average Price Path



Parameters: $T = 10$, $v = 100$, $\delta = 1$, $\mu = 10$, myopic consumers. Initial capacities under duopoly are $x_1 = 8$, $x_2 = 6$ and under monopoly $x_M = 14$.

As an example, consider Figure 3, which for $T = 10$ shows representative equilibrium price and capacity paths as well as average price paths under duopoly and monopoly, given initial capacities of $x_1 = 8$ and $x_2 = 6$ as well as $x_M = 14$ respectively. The smaller duopolist undercuts the larger duopolist, while a monopolist with excess capacities posts the final-period price in all time periods.[27]

[27]Note that the average price path (Figure 3 b) of the smaller firm 2 decrease towards the deadline because these prices reflect the average of rare path realizations where firm 2 will not have sold all its capacities by then. In the representative path (a) firm 2 sells its last capacity in $t = 3$.

Note that if the difference between firms' capacities is relatively large, then the high-capacity firm might post a lower price. Having only few capacities to sell off, the smaller firm then has the higher reservation value allowing it to post higher prices initially, as it can "bet" on some positive taste realizations. Contrary, the larger firm has more capacities to sell off and is therefore relatively more eager to sell, making this case similar to the capacity scarcity case discussed below. The example in Figure A.2 of Appendix A.3 shows this situation with initial capacities of $x_1 = 12$ and $x_2 = 2$ for $T = 10$.

Symmetric Aggregate Excess Capacities ($\sum_i x_i > t$, $x_i \approx x_j < t$)

If the sum of firms' capacities exceeds the number of remaining time periods and firms' capacities are (relatively) equal but neither firm can serve the entire expected market itself, competition can be extreme.[28] In the dynamic Bertrand model (Dudey, 1992) pricing below marginal costs can result. Similarly in my model, firms will compete harshly to become the smaller firm. Intuitively, firms anticipate that becoming the smaller firm subsequently allows them to sell their capacities earlier. Thus they fiercely compete down to a (possibly negative) price at which they become indifferent between selling now at this non-profitable price but then continue as the more profitable smaller firm, and the continuation value of remaining as the larger firm. However, this effect is damped by consumer heterogeneity, as further explained in Section 5.3.[29]

As an example, consider Figure 4, which for $T = 10$ shows representative equilibrium paths as well as average price paths, given initial capacities of $x_1 = 7$ and $x_2 = 7$ under duopoly as well as $x_M = 14$ under monopoly as a reference. Note that in the representative path (a) negative prices are posted in $t = 10$ when $x_1 = x_2 = 7$ and also in $t = 8$ when $x_1 = x_2 = 6$.

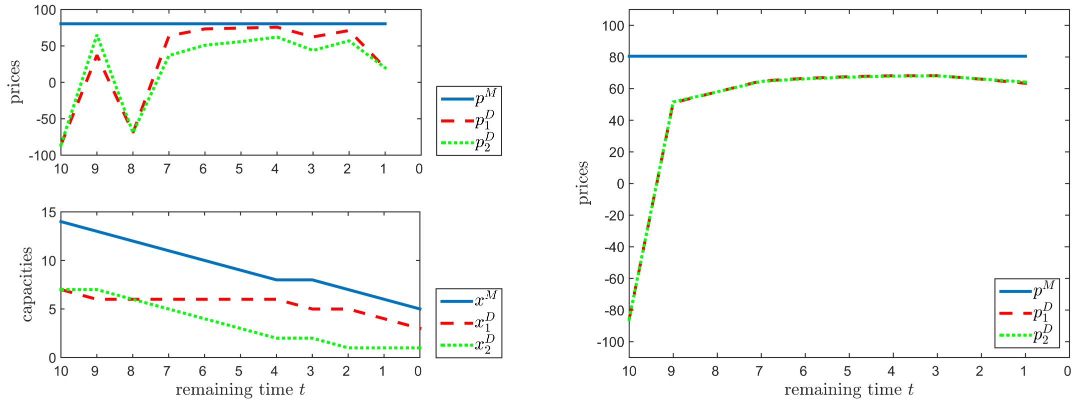
Result 4. *Under symmetric aggregate excess capacities there is harsh competition, even with negative prices, to become the smaller firm.*

[28] Note that due to consumer heterogeneity such a situation might also occur for other than but close to equal capacities. In Figure 2 this capacity case does nor occur at all.

[29] Note that under symmetric aggregate excess capacities many comparative static results will seem initially counter-intuitive due to this "race" to become the smaller firm.

Figure 4: Symmetric Aggregate Excess Capacities

(a) Representative Price and Capacity Path (b) Average Price Path



Parameters: $T = 10$, $v = 100$, $\delta = 1$, $\mu = 10$, myopic consumers. Initial capacities under duopoly are $x_1 = 7$, $x_2 = 7$ and under monopoly $x_M = 14$.

Scarce Capacities ($\sum_i x_i \leq t$)

If the sum of both firms' capacities is not larger than the number of expected consumers, i.e. capacities are scarce ($x_1 + x_2 = 6 \leq t$ in Figure 2), in static and dynamic Bertrand models both firms can sustain monopoly prices and they extract all consumer valuation throughout the entire selling horizon. In my model, consumer heterogeneity can actually reduce prices such that lower realizations of consumer taste can be accounted for. Further, the firm with higher remaining capacity undercuts the other firm in order to increase its selling probability.[30] This incentive arises because given the possibility of lower consumer taste realizations the larger firm faces a higher opportunity cost if no sale occurs, i.e. fears more the continuation in the asymmetric aggregate excess capacities case. Contrary, the smaller firm will not follow in this price reduction and the intuition is that here it has the higher reservation value because capacities are scarce.

A monopolist with scarce capacity has a higher reservation value and posts higher prices than with excess capacity. However, if she does not sell and hence the number of remaining time periods decreases further towards a level closer to her capacity, her reservation value falls and she lowers prices towards her static excess capacity price.

Consider the example in Figure 5, which for $T = 10$ shows representative equilibrium paths as well average price paths, given initial capacities of $x_1 = 4$ and $x_2 = 2$

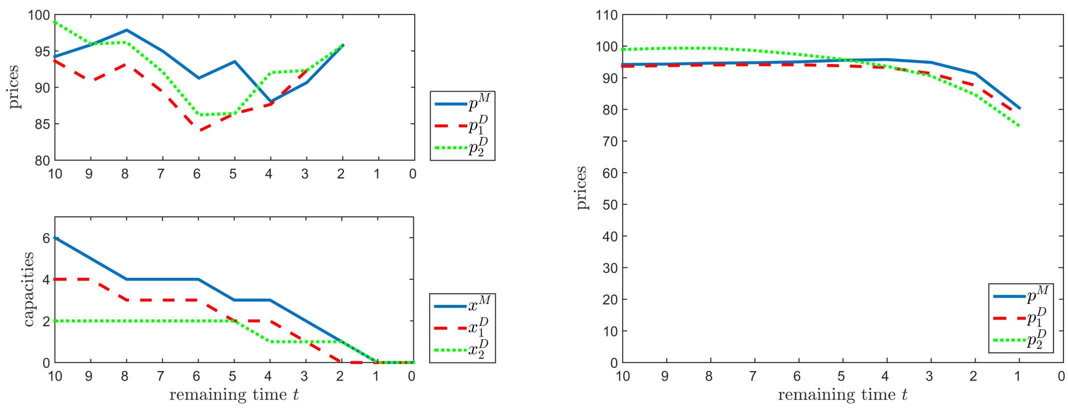
[30] Note that for $\sum_i x_i = t$ and large t, the smaller firm might slightly undercut, as this situation approaches the case of aggregate excess capacities.

under duopoly, while $x_M = 6$ under monopoly initially. Note that the monopoly price might even be below a duopolist's price but if the prices are weighted with their demand probability, the expected price of a sale under duopoly is typically still smaller.[31]

Result 5. *Under scarce capacities competitive prices are high but dispersed and the larger firm undercuts the smaller firm.*

FIGURE 5: Scarce Capacities

(a) Representative Price and Capacity Path (b) Average Price Path



Parameters: $T = 10$, $v = 100$, $\delta = 1$, $\mu = 10$, myopic consumers. Initial capacities under duopoly are $x_1 = 4$, $x_2 = 2$ and under monopoly $x_M = 6$.

5.2 Capacity Levels and Remaining Time

Given the four capacity cases introduced in Section 5.1 now I will consider the effect of variation in firms' capacity levels on pricing in a fixed period and also how prices change in remaining time for fixed levels of capacity. This provides interesting results on price monotonicity and additionally sets the groundwork for an assessment of the ex-ante game of capacity production (Section 6.4).

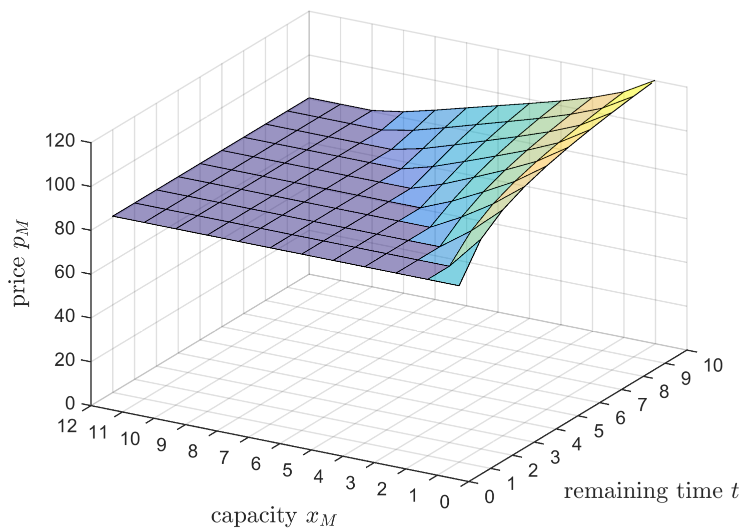
First, consider the monopoly situation. Without consumer heterogeneity, as in Dudey (1992), monopoly prices are always equal to valuation v. In my model pricing differs, as seen in Figure 6. Here monopoly prices (vertical axis) are plotted for different levels of fixed monopoly capacities and remaining time periods (horizontal axes). Immediately we can see that monopoly prices weakly increase in time and weakly decrease in capacity, ceteris paribus. Intuitively, given fixed capacities a monopolist's reservation value increases if she has more time to sell her goods. Contrary, in a given period her reservation value

[31] In Section 5.6, which is on market power, this is discussed in more detail.

is lower, the more capacity she holds. However, if capacity becomes excessive relative to remaining time ($x_M > t$), monopoly prices are constant and equal to the final-period price. Hence, the following result.[32]

Result 6. *Under monopoly, prices are weakly decreasing in capacity and weakly increasing in remaining time.*

FIGURE 6: Comparative Statics of Capacity and Time (Monopoly)



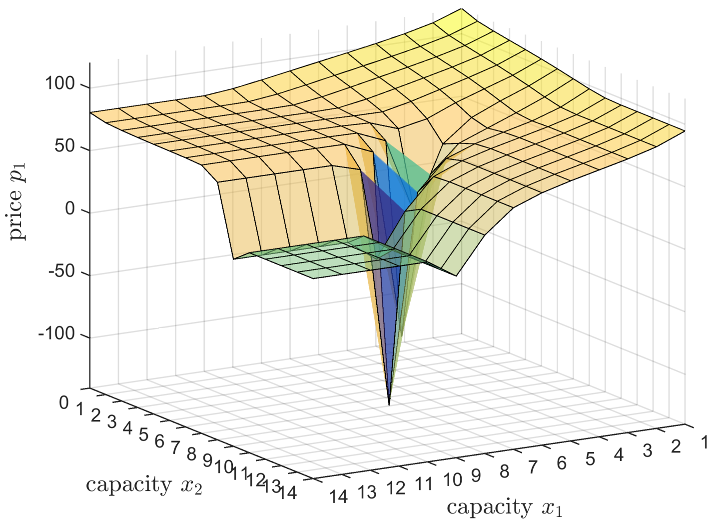
Parameters: $v = 100$, $\delta = 1$, $\mu = 10$, myopic consumers. Monopoly prices p_M for different capacity levels of x_M and different remaining time periods t.

Second, to see how duopoly prices in a given period change in capacity level variation we have to distinguish between the different capacity cases. For this consider the example in Figure 7 for $t = 10$ showing firm 1's price (vertical axis), given variation in its own and also in firm 2's capacity levels (horizontal axes). If there are individual excess capacities, i.e. each firm can serve the entire expected demand (in the figure $x_i \geq 10 = t$ for $i = 1, 2$), pricing is unaffected by the exact capacity levels because the period is essentially treated like the final one and firms post the static total-competition price (compare Proposition 3). Contrary, if there is at least one active firm that cannot serve all expected demand, i.e. $\min_i\{x_i\} < t$ and $i \in J(\boldsymbol{\omega})$, today's prices and hence selling probabilities do have implications for reservation values, i.e. future prices and selling probabilities. Specifically, if together the firms cannot serve the entire demand, i.e. capacities are scarce (in the

[32] This result is robust to forward-looking consumers, for this consider Figure A.3 in Appendix A.3.

figure if $x_1 + x_2 \leq 10 = t$), a firm's price decreases in its own and also in the competitor's capacity. Intuitively, lower own or lower competitor capacity and hence lower aggregate capacity means higher reservation values because firms are more confident to sell out eventually and hence can post higher prices.[33] However, if we have aggregate excess capacities and rather symmetric capacities (in the figure $x_1 \approx x_2 \in [6, 9]$), pricing can be extremely aggressive (compare Section 5.1), leading to possible non-monotonicity of prices in capacity levels, i.e. a steep decrease followed by an increase of prices in any firm's capacity.[34]

FIGURE 7: Comparative Statics of Capacity Levels (Duopoly)



Parameters: $t = 10$, $v = 100$, $\delta = 1$, $\mu = 10$, myopic consumers. Price p_1 in $t = 10$ under duopoly for different capacity levels of x_1 and x_2.

These observations prove the following findings, which are robust to forward-looking consumers, as illustrated in Figure A.4 of Appendix A.3.

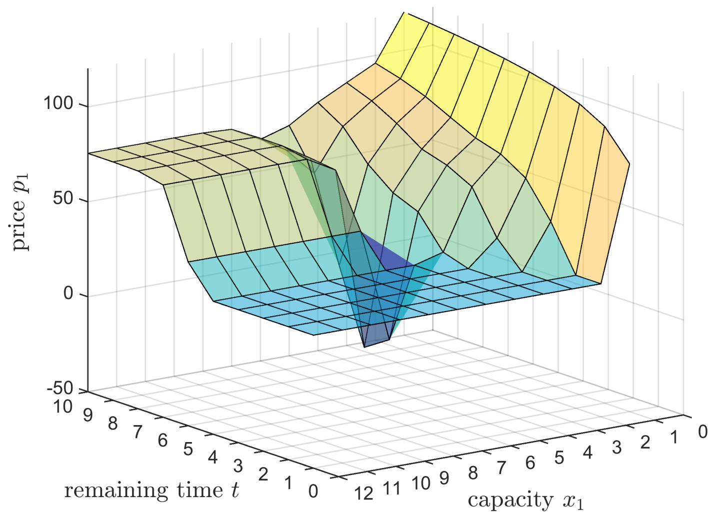
Proposition 7. *Under duopoly in a fixed period, prices are neither monotone in own nor in competitor capacity.*

Result 7. *If capacities are scarce in a fixed period, prices decrease in own and in competitor capacity.*

[33]Note for very low capacities that it might be that $p_i > v$, as the firms are comfortable to "bet" on high taste realizations as long as the deadline is still relatively far.

[34]Similarly, if $x_j > t$ but x_i is just lower than t, e.g. such that $x_1 = 9 < 10 = t < x_2$, firms' reservation values are relatively low and they are eager to sell because if no sale occurs, in the subsequent periods there will be individual excess capacities and hence total-competition pricing until the deadline.

FIGURE 8: Comparative Statics of Remaining Time (Duopoly)



Parameters: $v = 100$, $\delta = 1$, $\mu = 10$, myopic consumers. Price p_1 under duopoly for different levels of capacity x_1 and different remaining time periods t, while $x_2 = 5$ is fixed.

Finally, I consider how under duopoly a firm's price changes in remaining time t, for fixed levels of capacity. Figure 8 shows firm 1's price (vertical axis) for various fixed levels of capacity x_1 in remaining time t (horizontal axes), whereby competitor capacity is fixed at $x_2 = 5$. Similar to the analysis above we see that prices are constant in t and equal to the final-period total-competition price when capacities are individually excessive (here for $x_1 \geq 5 = x_2 \geq t$). Further, prices increase in remaining selling time if capacities are scarce (here if $x_1 + 5 \leq t$) because, intuitively, more time to sell their goods increases firms' reservation values. Only when capacities are on aggregate excessive, prices might not be monotone in remaining time. As discussed above, specifically if capacities are sufficiently symmetric (here for example $x_1 = x_2 = 5 < 6 = t$), prices become very aggressive (even negative) while they increase again to the total-competition price as soon as capacities become individually excessive relative to remaining time (here for example $x_1 = x_2 = 5 = t$). The intuition is that firms "race" to become the smaller firm (compare Section 5.1). Finally, note that without consumer heterogeneity (Dudey, 1992) prices can also be non-monotonic in capacity and remaining time, yet remain constant and equal to the monopoly price under scarcity. The following findings summarize these observations.[35]

[35] This is robust to forward-looking consumers. See Figure A.5 in Appendix A.3 and refer to the discussion in Section 6.2 for why under individual excess capacities prices decrease in remaining time.

Proposition 8. *Under duopoly with fixed capacities, a firm's price is not monotone in remaining time.*

Result 8. *If capacities are scarce and fixed, prices increase in remaining time.*

The analysis of this section shows that prices will typically increase after a sale and decrease otherwise. This feature is common with the literature of monopoly and oligopoly dynamic pricing models, for example in Talluri and Van Ryzin (2006), Hörner and Samuelson (2011) or Dilme and Li (2016). However, my model dynamics explain additional price volatility, too. As Propositions 7 and 8 suggest, under competition prices can go up or down in the next period, regardless of trade actually having taken place. Technically, this non-monotonicity emerges because any firm could sell (or not) in a given period due to consumer heterogeneity, such that different capacity cases with distinct pricing can occur.

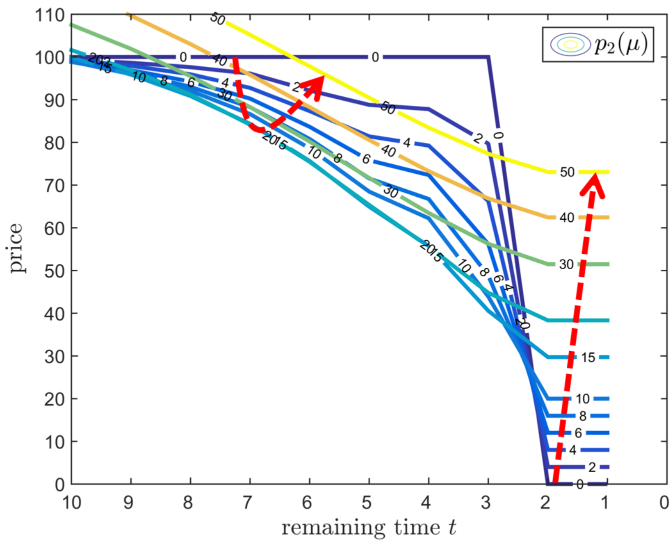
5.3 Consumer Heterogeneity

The (baseline) setting in this paper differs from Dudey's (1992) primarily in the introduction of consumer heterogeneity. For an assessment of the impact of consumer heterogeneity on equilibrium outcomes, remember that buying from a firm i gives a consumer utility $U_i = v - p_i + \epsilon_i \mu$. For $\mu \to 0$ the last term vanishes and we essentially have Bertrand demand.[36] Consequently, also the equilibrium prices must tend to their dynamic Bertrand equivalents.

To see this for an initial example consider the situation from before, where I showed equilibrium prices for fixed capacities $x_1 = 4$ and $x_2 = 2$, depending on remaining selling time. Figure 9 shows firm 2's price, for various levels of consumer heterogeneity μ. Note that as long as not every firm can serve all expected consumers, i.e. in the region $t > 2$, for $\mu \to 0$ we tend to monopoly prices equal to valuation v, but as soon as even the small firm's capacity becomes excessive relative to expected demand, i.e. in the region for $t \leq 2$, prices become competitive and jump down towards marginal costs. In the former region the price initially decreases and then increases in μ, while in the latter region prices increase in all μ, as indicated by the dashed arrows. To see that this is robust with forward-looking consumers, refer to Figure A.6 of Appendix A.3, which presents the respective example for forward-looking consumers.

[36] In Appendix A.1 I show that then the demand functions tend to their Bertrand equivalents.

FIGURE 9: Comparative Statics of Consumer Heterogeneity



Parameters: $T = 10$, $v = 100$, $\delta = 1$, myopic consumers. Capacities are fixed for the entire time horizon, s.t. $x_1 = 4$, $x_2 = 2$ in all t. The price $p_2(\mu)$ is for different levels of consumer heterogeneity $\mu \in \{0, 2, 4, 6, 8, 10, 15, 20, 30, 40, 50\}$, where the "in-line" number refers to μ. The values for $\mu = 0$ are from Dudey (1992). The dashed arrows show the effect of μ on p_2, one for the region $x_2 = 2 < t$ and the other for the region $x_2 = 2 \geq t$.

Intuitively, two effects of consumer herterogeneity are relevant here. On the one hand, heterogeneity essentially increases firms' market power as goods become less homogeneous, allowing firms to increase prices. On the other hand, heterogeneity increases the relative value of the outside option and hence firms react by decreasing prices. Which effects dominates depends on the capacity situation and the level of μ. The following result summarizes the section's findings on this, while its details will be discussed in the subsequent case analysis.

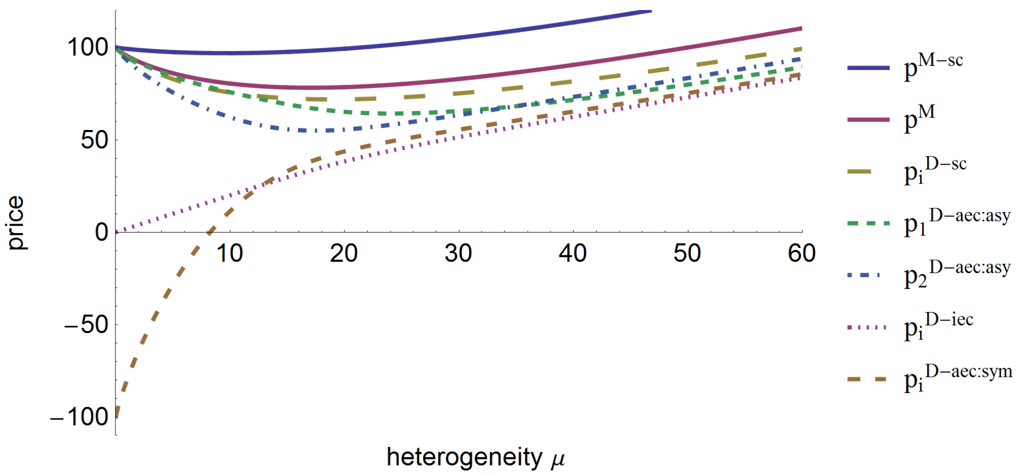
Result 9. *If the market is very competitive, i.e. under individual excess capacities or symmetric aggregate capacities, prices increase in consumer heterogeneity μ. Under all other capacity distributions as well as under under monopoly prices decrease for small μ and increase for larger μ.*

Individual Excess Capacities $(\min_i\{x_i\} \geq t,\ i \in J(\boldsymbol{\omega}))$
or Symmetric Aggregate Excess Capacities $(\sum_i x_i > t,\ x_i \approx x_j < t)$

If each firm can serve individually all expected consumers or firms have symmetric but on aggregate excess capacities, prices increase in consumer heterogeneity, $\partial p_i^*/\partial\mu > 0$ for all μ. Intuitively, in both these cases competition is relatively strong such that the first effect of increased market-power through heterogeneity is relatively more important than the second outside-option effect for all μ.

For the individual excess capacity case note that for $\mu \to 0$ we get $p_i^* \to 0$. To see that the price for an increased μ must also increase, consider the equilibrium value function given in equation (13) and rearrange to $p_i^*(\boldsymbol{\omega}) = \mu + V_i(\boldsymbol{p}^*, \boldsymbol{\omega}) - \delta W_{i,i}(\boldsymbol{\omega})$. Note then that here $p_i^*(\boldsymbol{\omega}) \geq \mu$ because $V_i(\boldsymbol{p}^*, \boldsymbol{\omega}) - \delta W_{i,i}(\boldsymbol{\omega}) \geq 0$, since under individual excess capacities the continuation value after a successful sale cannot exceed the total value function. Therefore, the price $p_i^*(\boldsymbol{\omega})$ must be at least as large as μ and hence if $\mu > 0$, also $p_i^* > 0$. This only shows that prices are larger than zero for any $\mu > 0$. Figure 10 illustrates that prices are increasing in μ for all μ with individual excess capacities under duopoly (labeled p_i^{D-iec}).

FIGURE 10: Comparative Statics of Consumer Heterogeneity (2)



Parameters: $t = 4$, $v = 100$, $\delta = 1$, myopic consumers
- Monopoly (M-sc) scarce capacity: $x^{M-sc} = 2$
- Monopoly (M) capacity: $x^M \geq 4$
- Scarce capacities duopoly (D-sc): $x_i^{D-sc} = 2$, $i \in \{1, 2\}$
- Asymmetric aggregate excess capacities duopoly (D-aec:asy): $x_1^{D-asc.asy} = 2$ and $x_2^{D-asc} = 4$
- Individual excess capacities duopoly (D-iec:asy): $x_i^{D-iec} = 4$, $i \in \{1, 2\}$
- Symmetric aggregate excess capacities duopoly (D-aec:sym): $x_i^{D-asc:sym} = 3$

For the symmetric aggregate excess capacity case (labeled $p_i^{D-aec:sym}$) note that for $\mu \to 0$ in the duopoly equilibrium we have $p_i^* \to v(t+1-2x_i)$, i.e. non-positive prices (Dudey, 1992). With $\mu > 0$ firms will also compete harshly to become the smaller firm, however this is dampened by consumer heterogeneity, which allows for more random demand later, such that being the smaller firm does not guarantee selling out first.

Asymmetric Aggregate Excess Capacities ($\sum_i x_i > t \wedge \min_i\{x_i\} < t,\ x_i \neq x_j$) or Scarce Capacities ($\sum_i x_i \leq t$)

If there is a duopoly with asymmetric aggregate excess capacities (labeled $p_i^{D-aec:asy}$ in Figure 10) or scarce capacities (p_i^{D-sc}), and also if there is a monopoly with excess capacity (p^M) or with scarce capacity (p^{M-sc}), then without consumer heterogeneity, i.e. for $\mu \to 0$, prices tend to the monopoly price v. For $\mu > 0$ simulations suggest that prices fall from v, such that $\partial p_i^*/\partial\mu < 0$, but only as long as μ is not too large, while we get $\partial p_i^*/\partial\mu > 0$ if μ is large.[37] Intuitively, in all these capacity cases both the market-power and the outside-option effect of consumer heterogeneity are relevant. For smaller μ the latter effect dominates, while for higher μ the market-power effect becomes increasingly more important.[38]

5.4 Consumers Patience

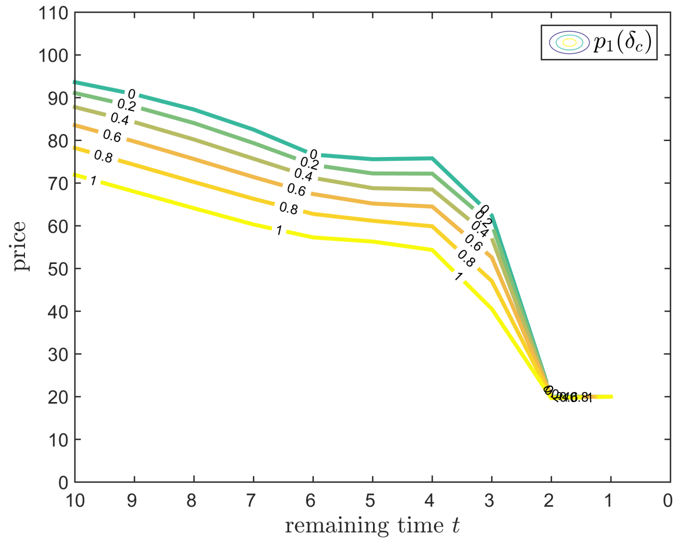
In this section I will discuss the impact of consumer patience on equilibrium prices. For this I make use of the extended model with forward-looking consumers as presented in Section 4, i.e. consumers can choose to wait in the market instead of exiting the market. The parameter defining consumers' patience level is δ_c, the discount factor of consumers. Fully patient consumers do not discount the future at all (not more than firms), such that $\delta_c = \delta$, whereas completely impatient consumers disregard future payoffs, such that $\delta_c = 0$.[39] Note that with completely impatient consumers we are back in the baseline model of myopic consumers.

[37] Note that if capacities are very scarce, prices still fall for small μ but start increasing quickly in μ, s.t. they can reach values above v already for relatively small μ.

[38] The multinomial-logit demand becomes essentially random for very high values of consumer heterogeneity and therefore increasingly inelastic. For example, if $v = 100$, for the monopoly price we get $\partial p_M^*/\partial\mu > 1$ for all $\mu > 60$ and therefore prices increasingly larger than v. Similarly, this is true under duopoly. Therefore, it seems sensible to restrict μ and focus on values of μ such that at least the total-competition price p_i^* as defined in equation (15) is below consumer valuation v. For $v = 100$ this is true for all $\mu < 60$.

[39] Technically, also $\delta_c = 1 < \delta$ is possible, however not sensible.

Figure 11: Comparative Statics of Consumer Patience



Parameters: $T = 10$, $v = 100$, $\delta = 1$, $\mu = 10$, forward-looking consumers. Capacities are fixed for the entire time horizon, s.t. $x_1 = 4$, $x_2 = 2$ for all t. Price p_1 under duopoly for different levels of consumer discount levels $\delta_c \in \{0, 0.2, 0.4, 0.6, 0.8, 1)$, where the "in-line" number refers to δ_c.

At first consider the example from before in Figure 11, where I depict firm 1's equilibrium price for fixed capacities, given different values of δ_c. Note that prices are generally lower, the more patient consumers are, while they remain constant in δ_c under individual excess capacities, i.e. as soon as even $x_2 = 2 \geq t$.

The following lemma provides an analytical outcome for $T = 2$, while for higher T I provide simulation results below.

Lemma 3. *For $T = 2$, prices weakly decrease if consumers become more patient, i.e. more forward-looking.*

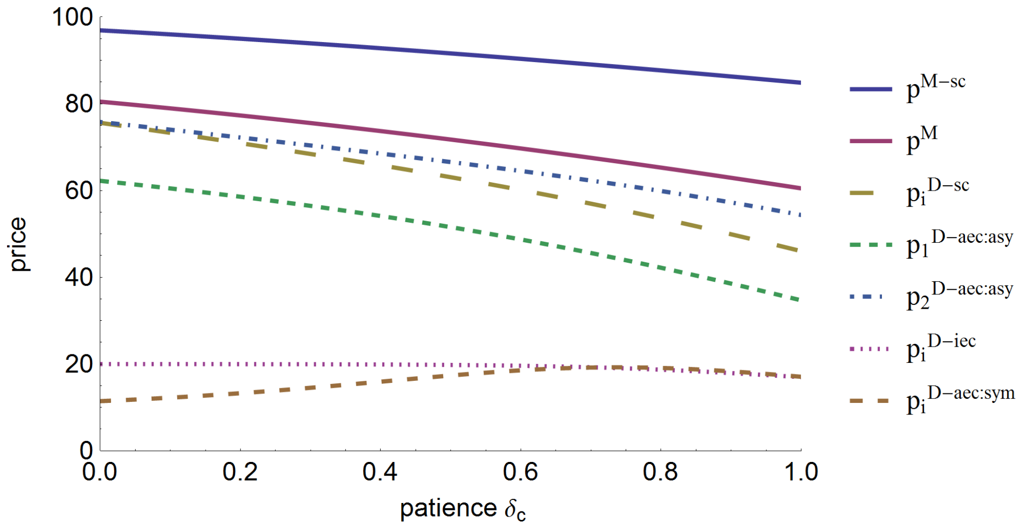
Although more complex due to δ_c's impact on $W_c(\boldsymbol{\omega})$ as well as on $W_{i,i}(\boldsymbol{\omega}) - W_{i,j}(\boldsymbol{\omega})$, simulation results suggest that for higher T the effect is similar.

Result 10. *Prices weakly decrease if consumers become more patient, i.e. more forward-looking, unless there are symmetric aggregate excess capacities.*

The intuition is straightforward. The more patient consumers are, the higher their valuation of waiting, such that demand for firms' goods decreases. Consequently firms will decrease prices.

For $t = 4$ Figure 12 illustrates that prices weakly decrease for all levels of δ_c in all distinguished capacity cases. Only in the case of symmetric aggregate excess capacities (compare Section 5.3), prices (initially) increase in δ_c. Intuitively, if consumers become more patient, they choose to wait more often, hence firms anticipate that they might not remain in the aggregate excess capacity case but could move into the individually excess capacity case, which would lead to generally lower prices. Therefore, the continuation value of becoming the smaller firm decreases and firms would not be willing to accept such a low (or negative) initial price, hence prices slightly increase. If consumers become even more patient, pressure on demand (and hence prices) becomes higher such that all prices decrease in equilibrium, even under symmetric aggregate excessive capacities.

FIGURE 12: Comparative Statics of Consumer Patience (2)



Parameters: $t = 4$, $v = 100$, $\delta = 1$, $\mu = 10$
- Monopoly (M-sc) scarce capacity: $x^{M-sc} = 2$
- Monopoly (M) capacity: $x^M \geq 4$
- Scarce capacities duopoly (D-sc): $x_i^{D-sc} = 2$, $i \in \{1, 2\}$
- Asymmetric aggregate excess capacities duopoly (D-aec:asy): $x_1^{D-aec:asy} = 2$, $x_2^{D-aec:asy} = 4$
- Individual excess capacities duopoly (D-iec:asy): $x_i^{D-iec} = 4$, $i \in \{1, 2\}$
- Symmetric aggregate excess capacities duopoly (D-aec:sym): $x_i^{D-asc:sym} = 3$

5.5 Consumer Arrival

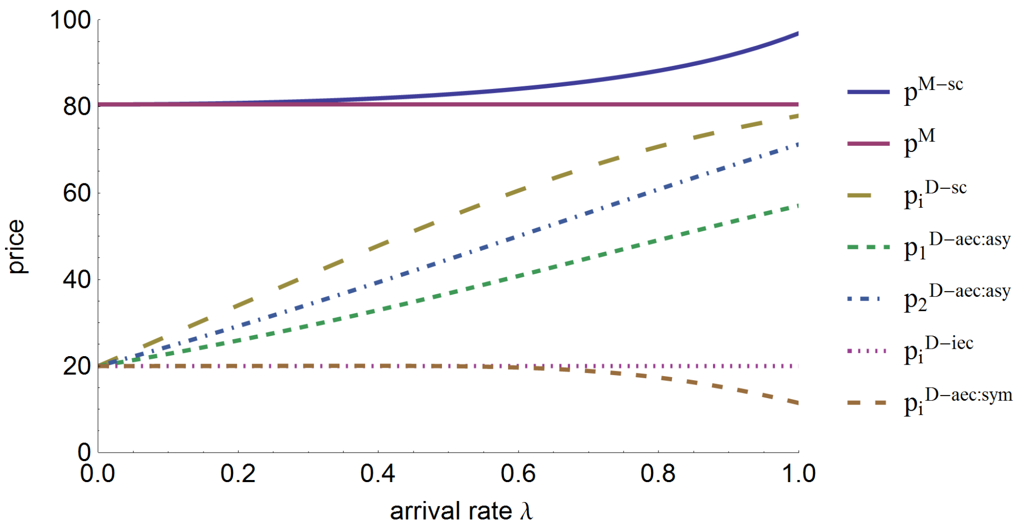
In the baseline model in every time period one additional consumer arrives with probability one. Now consider a general arrival rate, such that a new consumer only arrives with probability $\lambda \leq 1$ each period. Let the arrival rate λ be common knowledge. Remember that the timing in a period is such that first consumer arrival realizes, then all firms learn the number of consumers in the market and simultaneously post prices. Intuitively, if firms expect lower consumer arrival rates, expectations for future demand decrease, such that firms' reservation values decrease as they become less "confident" to sell out their goods, and hence post lower prices. For $T = 2$ consider the following lemma, while for higher T I provide simulation results below.

Lemma 4. *For $T = 2$, prices weakly decrease if the consumer arrival rate falls.*

For $T > 2$ the argumentation is more complex because the sign of $\delta W_{i,i}(\boldsymbol{\omega}) - \delta W_{i,j}(\boldsymbol{\omega})$ becomes ambiguous, especially in the case of symmetric aggregate excess capacities, however note the following simulation result.[40]

Result 11. *Prices weakly decrease if the consumer arrival rate falls, unless there are symmetric aggregate excess capacities.*

FIGURE 13: Comparative Statics of Consumer Arrival Rate



Parameters: $t = 4$, $v = 100$, $\delta = 1$, $\mu = 10$, myopic consumers
- Monopoly (M) scarce capacity: $x^{M-sc} = 2$
- Monopoly (M) capacity: $x^{M} \geq 4$
- Scarce capacities duopoly (D-sc): $x_i^{D-sc} = 2$, $i \in \{1, 2\}$
- Asymmetric aggregate excess capacities duopoly (D-aec:asy): $x_1^{D-aec:asy} = 2$, $x_2^{D-aec:as} = 4$
- Individual excess capacities duopoly (D-iec:asy): $x_i^{D-iec} = 4$, $i \in \{1, 2\}$
- Symmetric aggregate excess capacities duopoly (D-aec:sym): $x_i^{D-asc:sym} = 3$

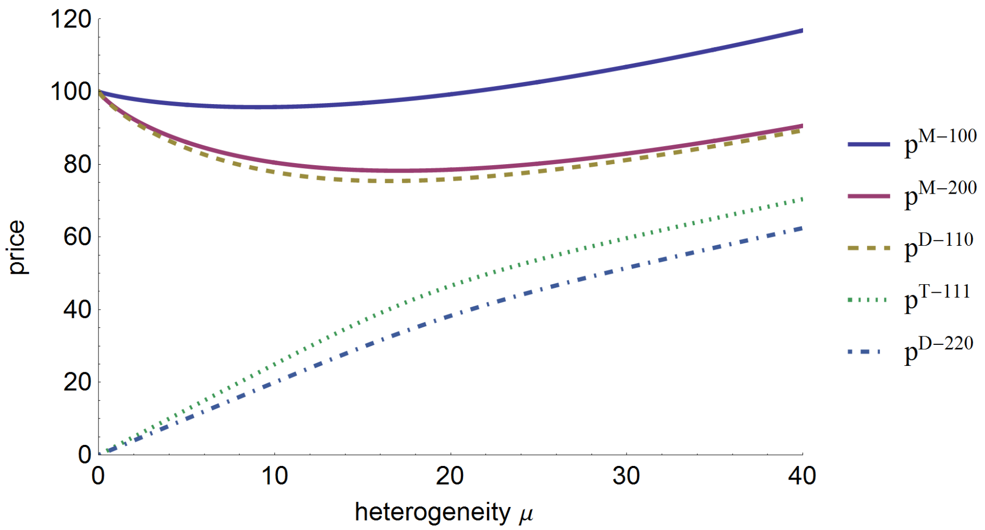
[40] When consumers are forward-looking, the same intuition applies and hence prices should increase in the arrival rate, too.

Consider Figure 13 providing an example for $t = 4$ to see that prices increase in the arrival rate under all capacity cases. Only if there are symmetric aggregate excess capacities, i.e. possibly below-marginal-cost pricing, prices might decrease for higher levels of the arrival rate λ. Similarly to the intuition for lower consumer patience δ_c, if the arrival rate of consumers is expected to be high and hence future demand to be more stable, symmetric firms compete even fiercer to become the smaller firm. However, if firms do not expect high consumer arrival rates, the continuation value of becoming the smaller firm decreases.

5.6 Market Power

To assess the effect of market power on prices I compare prices under monopoly and competition. At the same time I want to show that these comparisons are robust to different levels of consumer heterogeneity μ. For $\mu \to 0$, the monopoly price is always equal to valuation v, while competitive prices are also equal to v, unless there are individual excess capacities, which yield prices equal to marginal costs.[41]

FIGURE 14: Monopoly vs Competition ($t = 2$)

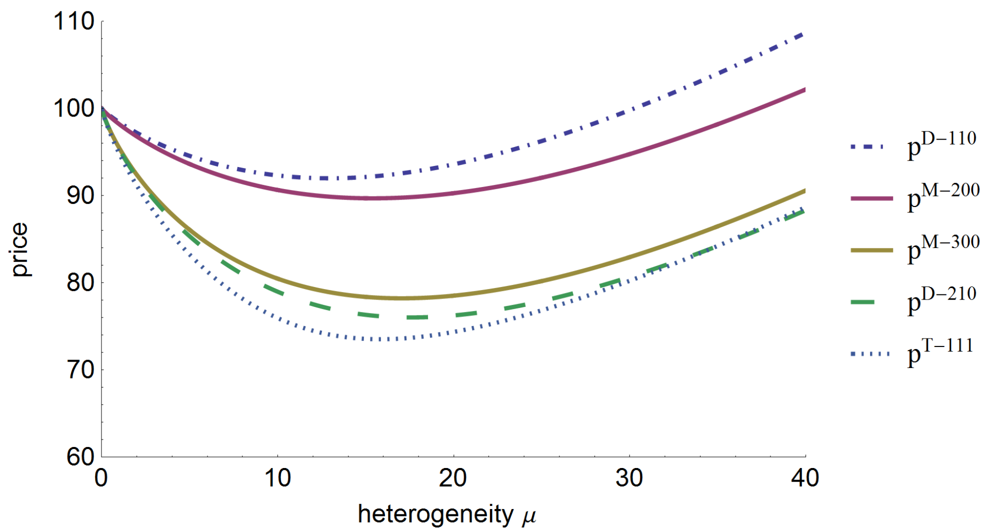


Parameters: $t = 2$, $v = 100$, $\delta = 1$, myopic consumers
- Monopoly (M) capacity: $x^{M-100} = 1$, $x^{M-200} \geq 2$
- Duopoly (D-110): $x_1^{D-110} = x_2^{D-110} = 1$, $x_3^{D-110} = 0$
- Duopoly (D-220): $x_1^{D-220} = x_2^{D-220} = 2$, $x_3^{D-220} = 0$
- Triopoly (T): $x_i^{T-111} = 1$, $i \in \{1, 2, 3\}$

[41] Only symmetric aggregate excess capacities yield prices below marginal costs.

Figure 14 provides an example for $t = 2$, where for $\mu > 0$ monopoly and duopoly prices differ, however remain “similar” when the aggregate market capacities are equal and scarce.[42] In a market with three firms (triopoly) or aggregate excess capacities, prices are substantially lower (and equal to marginal costs for $\mu \to 0$). This effect is standard: more firms and/or more capacity mean more aggressive pricing as the continuation value after no sale decreases.

FIGURE 15: Monopoly vs Competition ($t = 3$)



Parameters: $t = 3$, $v = 100$, $\delta = 1$, myopic consumers
- Monopoly (M) capacities are s.t. $x^{M-200} = 2$, $x^{M-300} \geq 3$
- Duopoly (D-210) capacities are s.t. $x_1^{D-210} = 2$, $x_2^{D-210} = 1$, $x_3^{D-210} = 0$
- Duopoly (D-110) capacities are s.t. $x_1^{D-110} = x_2^{D-110} = 1$, $x_3^{D-110} = 0$
- Triopoly (T) capacities are $x_i^{D-111} = 1$, $i \in \{1, 2, 3\}$

However, it is not necessary the case that prices always decrease in the number of firms, if overall market capacities are equal and relatively scarce (and in some cases of extremely asymmetric aggregate excess capacities, too). For this consider the example illustrated in Figure 15 for $t = 3$. For $\sum_i x_i = 2$, duopoly prices (labeled p^{D-110}) are higher than in the corresponding monopoly (p^{M-200}). Here for $\mu > 0$ goods become less homogeneous such that a "Cournot" effect becomes relevant, which dominates the standard market-power effect: Under competition firms do not internalize the positive externality a low price imposes on the other firms. While a monopolist posts prices with the intent of selling both of her units eventually, competitive firms care only about selling

[42]Further details on this similarity are in Section 6.3.

their own capacity and do not internalize that no sale yields harsher competition in the next period for the other firm, too.[43]

Result 12. *Unless capacities are relatively scarce, prices typically decrease in the number of firms.*

Note however that if capacities are not too scarce, e.g. $\sum_i x_i = t = 3$ in Figure 15, prices still decrease in the number of firms ($p^{D-210} < p^{M-300}$). Even if for asymmetric duopolies the smaller firm's price might be above the monopoly price, nevertheless if duopoly prices are weighted with their corresponding demand, the expected duopoly price at which consumers buy (p^{D-210} in the figure) is smaller than the monopoly price.

6 Welfare and Policy

In this section I will discuss policy effects on welfare. Fist, I will consider the effect of having forward-looking rather than myopic consumers. Second, I will further investigate on market power and competition policy, while third I will study ex-ante capacity production.

6.1 Welfare Definitions

Before analyzing policies we need to define total sales, average prices, firm profits and consumer surplus.

Given a state $\boldsymbol{\omega}$, **total sales** are

$$S(\boldsymbol{\omega}) = \sum_{J(\boldsymbol{\omega})} D_j\Big(\mathbf{p}^*(\boldsymbol{\omega}), \boldsymbol{\omega}\Big) + \delta \sum_j \Bigg\{ D_j\Big(\mathbf{p}^*(\boldsymbol{\omega}), \boldsymbol{\omega}\Big) S\Big(\boldsymbol{\omega'}|(\boldsymbol{\omega}, j)\Big) \Bigg\},$$

where $j \in J(\boldsymbol{\omega}) \cup \{0\}$ and at the deadline $S(0, \cdot) = 0$. Total sales are defined as the equilibrium-path-probability-weighted discounted sum of probabilities that trade will happen in the given periods. I will interpret total sales as a proxy measure of **efficiency**. Every time a capacity unit is sold this creates a value, which through the price is split between the firm and the consumer.[44] I argue that total welfare can be approximated by the total number of these value creations.[45]

[43]To see that these results are similar for forward-looking consumers, refer to Section 6.3 where I discuss competition policy under both regimes.

[44]In Section 6.4 I consider capacity production costs.

[45]Note that summation of profits and consumer surplus is not sensible in the multinomial logit models as the overall scale of utility is only meaningful as a relative measure.

Given a state $\boldsymbol{\omega}$, **average prices** are

$$p(\boldsymbol{\omega}) = \frac{\sum_{J(\boldsymbol{\omega})} \left\{ D_j\Big(\mathbf{p}^*(\boldsymbol{\omega}), \boldsymbol{\omega}\Big)\, p_j^*(\boldsymbol{\omega}) \right\}}{\sum_{J(\boldsymbol{\omega})} D_j\Big(\mathbf{p}^*(\boldsymbol{\omega}), \boldsymbol{\omega}\Big)} + \sum_j \left\{ D_j\Big(\mathbf{p}^*(\boldsymbol{\omega}), \boldsymbol{\omega}\Big) p\Big(\boldsymbol{\omega'} | (\boldsymbol{\omega}, j)\Big) \right\},$$

while at the deadline no prices are posted. Average prices are the equilibrium-path-probability-weighted sum of expected prices posted in the given periods. In a period the expected price is obtained by weighting prices with their respective demand probabilities.

Given a state $\boldsymbol{\omega}$, a firm i's **profit** and total **industry profits** are

$$\Pi_i(\boldsymbol{\omega}) = V_i\Big(\boldsymbol{p}^*(\boldsymbol{\omega}), \boldsymbol{\omega}\Big) \quad \text{and} \quad \Pi(\boldsymbol{\omega}) = \sum_i \Pi_i(\boldsymbol{\omega}).$$

A firm's (expected) profit in a state is equal to its expected valuation function, i.e. its equilibrium-path-probability-weighted discounted sum of expected revenues. The sum of all firms' profits shall be called industry profits.

Given a state $\boldsymbol{\omega}$, **consumer surplus** is

$$CS(\boldsymbol{\omega}) = \ln\left[1 + \frac{1}{|J(\boldsymbol{\omega})|} \sum_{J(\boldsymbol{\omega})} \exp\left(\frac{v - p_j^*(\boldsymbol{\omega})}{\mu}\right)\right] + \delta \sum_j \left\{ D_j\Big(\mathbf{p}^*(\boldsymbol{\omega}), \boldsymbol{\omega}\Big) CS\Big(\boldsymbol{\omega'} | (\boldsymbol{\omega}, j)\Big) \right\}$$

where at the deadline $CS(0, \cdot) = 0$. Consumer surplus here is based on the standard definition for the logit model as e.g. in Train (2009), which is the logarithm of the sum of utility terms from products of the active firms and the outside option. Since the problem is of multiple periods, I add over all periods weighted with the equilibrium-path-probabilities. Note that this notion of consumer surplus is different than the definition of the consumer value function I used in equation (22). In (22) the valuation of a consumer sums over all states she might reach for the first time again as the lucky consumer, providing the foundation for her decision criterion of whether to wait or not. Here, we need the consumer surplus which is created whenever a good is sold. Therefore, choosing to leave the market in the myopic model or choosing to wait in the forward-looking model both give the consumer the normalized outside option utility for that particular period, such that consumer surplus under both regimes is comparable.[46] Further note that the

[46] If similar to (22) consumer surplus with forward-looking consumers included instead her expected continuation value after waiting, some sales would essentially be counted twice, as my definition of consumer surplus already sums over all states of the equilibrium path.

utility term from a firm i's product is multiplied by $1/|J(\boldsymbol{\omega})|$, i.e. the inverse number of active firms, such that consumers obtain the same surplus at equal prices independent of the number of firms. Otherwise the mere number of active firms, i.e. the number of choices in set $J(\boldsymbol{\omega})$, would increase consumer surplus, ceteris paribus. Since consumer surplus is only a relative measure, this normalization is without loss of generality.[47]

Note for all welfare measures defined here that future states are discounted with the common (firm) factor δ, which I assume to be equal to the planner's discount factor.[48] Also consumer surplus is discounted with δ because it measures the sum of current and future consumers' surpluses, which is relevant for the planner and not for the individual consumer when choosing whether to wait or to buy.

6.2 Forward-looking Consumers

Myopic consumers only consider the current period. Contrary, forward-looking consumers can anticipate all future prices and might prefer to wait whenever they expect better prices and a relatively low rationing risk. For this they need to compute all possible continuations, which requires knowledge of the current state.

This could be achieved through a policy informing consumers about the current state, e.g. by regulation enforcing firms to publicly announce their real-time capacity levels. In the markets for airfare tickets or other travel markets consumers are typically not informed about remaining capacity, while in other markets, such as theater or event ticket markets, this information tends to be publicly available through, e.g., seat maps. Another, perhaps more subtle, regulation could aim at increasing the degree to which consumers are forward-looking. E.g., by providing additional information on typical price paths or rationing risks, but without any capacity level disclosure requirement for firms, consumers could learn to better deduct information about firms' capacity levels from announced prices. This way, they could become more patient, such that they could compute continuation values more effectively. If forward-looking consumers (or total welfare) are better off than under myopia, regulators could consider such policies.

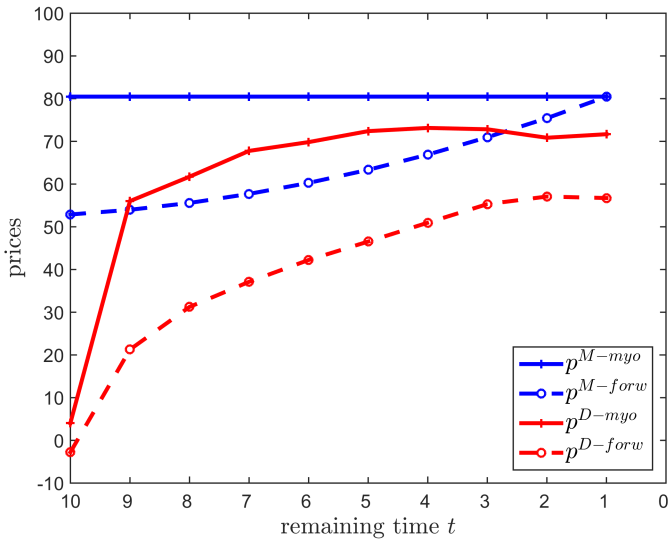
Technically, the myopic-consumer and the forward-looking-consumer models can be compared as the following three conditions hold in both models. First, in each time period at most one seller can sell a good. Second, if we start from the same initial state with no previous consumers in the market, then the total number of units that can maximally be sold until the deadline is identical in both models. This is because the

[47] Compare footnote 22 from the section on comparative static simulations for a discussion.

[48] This could be normalized to $\delta = 1$.

total number of consumers throughout the entire selling time is equal, even if consumers choose to wait for one or more periods. Third, the definitions of welfare measures as in Section 6.1 are identical under both regimes.

FIGURE 16: Forward-looking vs Myopic Consumers: Average Price Path



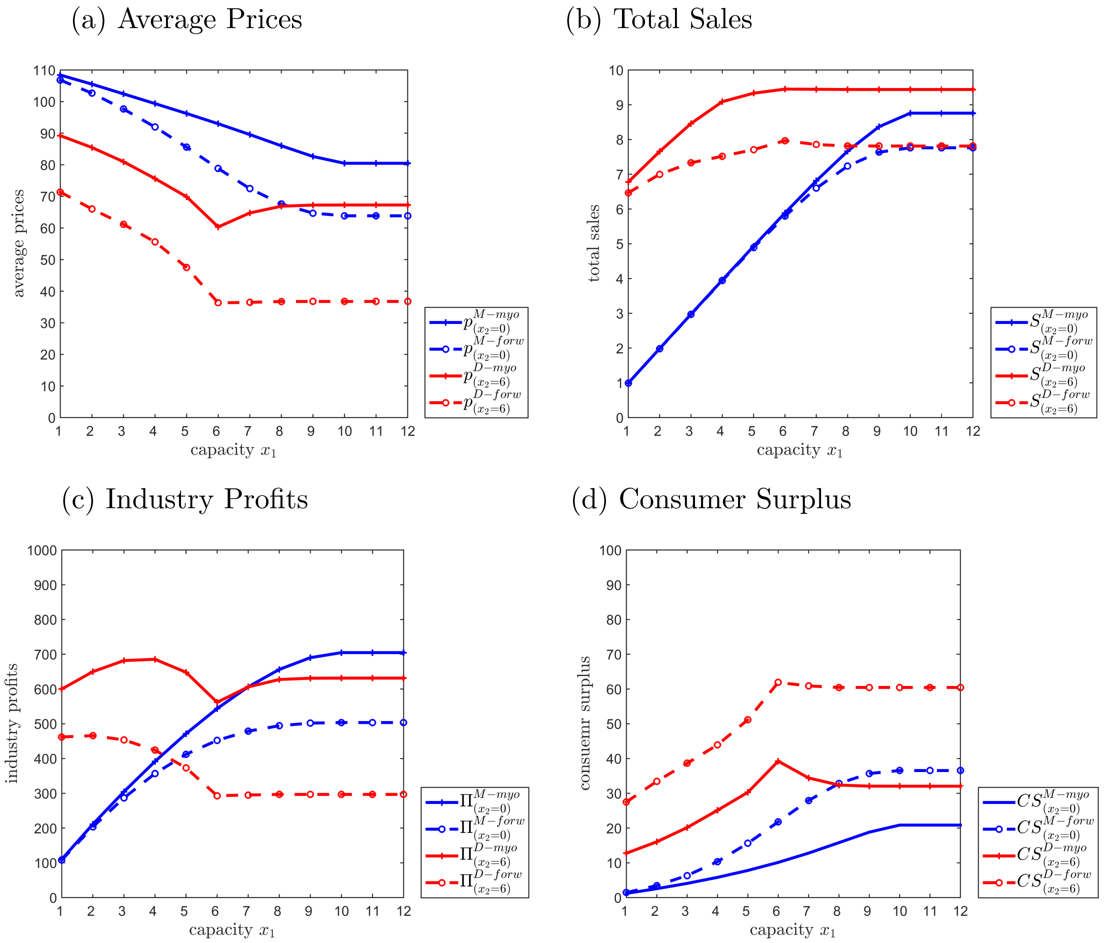
Parameters : $T = 10$, $v = 100$, $\delta = 1$, $\mu = 10$. Initial capacities at time $t = 10$ are given by $x_1 = 6$, $x_2 = 6$ under duopoly and by $x_M \geq 10$ under monopoly. Duopoly price p^D and monopoly price p^M, for myopic (myo: $\delta_c = 0$) and forward-looking consumers (forw: $\delta_c = \delta$).

From the comparative statics results from Section 5.4, we know that in a given state prices are lower when consumers are forward-looking. Therefore the average price path with forward-looking consumers entails lower prices than the respective price path with myopic consumers.[49] Consider the example in Figure 16, where for initially symmetric but aggregate excess capacities I show the average price path under duopoly and monopoly for myopic and completely forward-looking consumers. For the intuition of why prices with forward-looking consumers are generally lower, consider the monopoly price paths (p^{M-myo} and p^{M-forw}). Unlike in the case of myopic consumers, with forward-looking consumers the monopoly price is not constantly equal to the final-period price.

[49]Note that in the case of symmetric aggregate excess capacities and small levels of consumer patience δ_c prices can increase in consumer patience. However here we consider fully patient consumers where this is not the case. Nevertheless, even if the price in a state of symmetric aggregate capacity was higher for forward-looking consumers, this would yield lower trade probabilities and hence subsequently lower prices such that the average price (path) effect of forward-looking consumers should overall still be negative.

Although she holds excess capacities throughout the entire selling period, here a monopolist gradually increases prices. The longer the time until the deadline, the higher consumers' valuation for waiting as the rationing risk is still relatively small. Consequently, the posted prices must be lower if the deadline is still further away, while only in $t = 1$ the monopoly prices are identical in both cases. Respectively, this intuition carries over to competition.

FIGURE 17: Forward-looking vs Myopic Consumers: Welfare



Parameters: $t = 10$, $v = 100$, $\delta = 1$, $\mu = 10$. Average prices (a), total sales (b), industry profits (c) and consumer surplus (d), as expected in $t = 10$, under monopoly and duopoly for different levels of initial capacity x_1, given fixed levels for $x_2 \in \{0, 6\}$, for myopic (myo: $\delta_c = 0$) as well as forward-looking consumers (forw: $\delta_c = \delta$).

For the comparison of the previously introduced welfare measures under myopic and forward-looking consumers consider the example in Figure 17, where I plot these measures for different capacity levels of firm 1, as expected in period $t = 10$. Average prices (a) with myopic consumers are higher than with forward-looking consumers, under monopoly as well as under duopoly. Interestingly, although less goods are sold overall (b), prices under consumer myopia are sufficiently higher such that firms are still better off (c). Contrary, consumers (d) remain worse off when myopic. While here I provide the case where one firm cannot serve all expected demand, in Figure A.7 of Appendix A.3 I show that this is robust to the case where both firms can have excess capacities. In general, I find this result to be robust when looping over a wide range of parameter constellations in the simulations.

Result 13. *Average prices, total sales and industry profits are higher with myopic consumers than with forward-looking consumers, while consumer surplus is lower.*

For the regulator it is interesting that the efficiency loss of having forward-looking consumers (less total sales due to the waiting possibility) does not go in hand with a consumer surplus reduction because the threat of waiting sufficiently depresses prices.[50]

6.3 Competition Policy

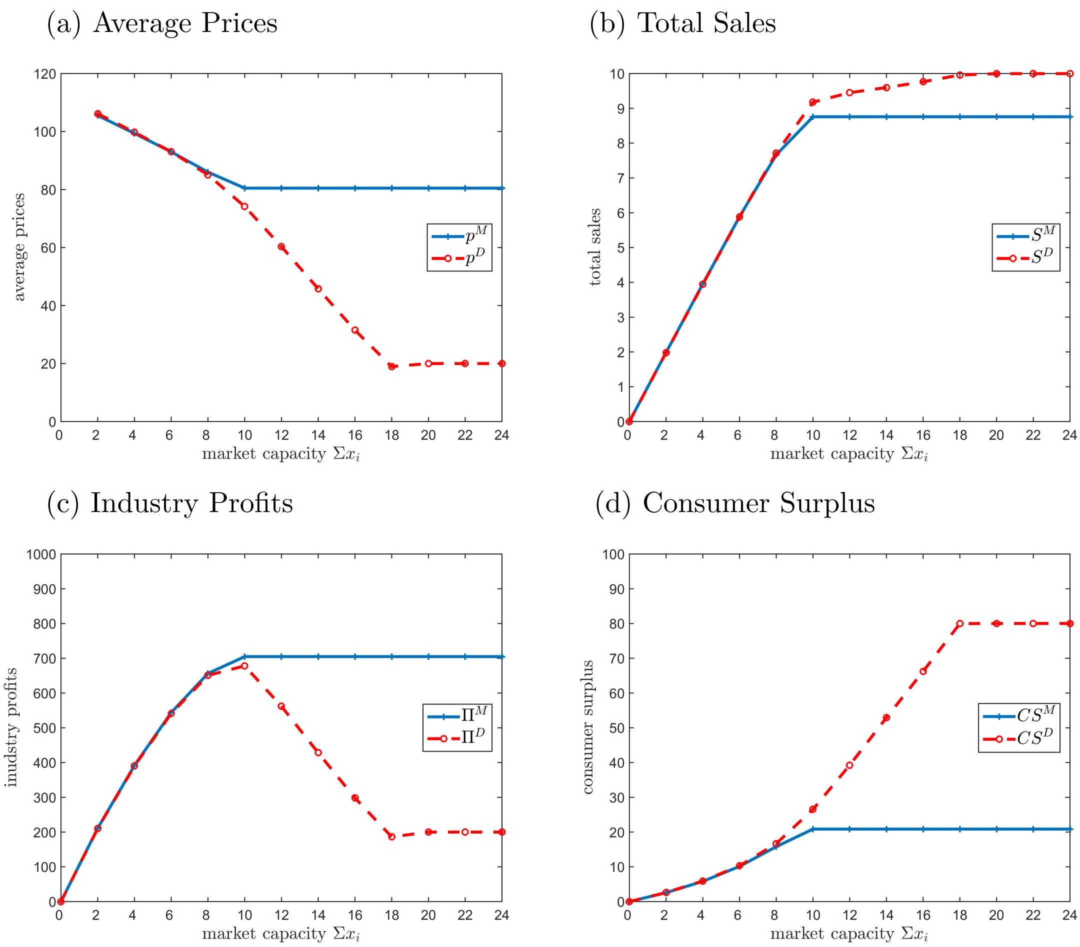
Competition policy and merger regulation in markets with dynamic pricing are omnipresent, e.g. in the assessment of airline mergers or booking platform market power. In this section I will analyze the effects of competition on welfare measures by comparing monopoly to duopoly situations.

While in chapter 5.6 I discussed how prices change in the number of firms, in Figure 18 I plot average prices, total sales, industry profits and consumer surplus, as expected in period $t = 10$, for different levels of total market capacity under monopoly and duopoly, whereby both duopolists initially hold exactly half the number of capacities as the monopolist. As long as we are under scarcity, i.e. whenever total market capacity is lower than the number of remaining selling periods such that $\sum_i x_i < t$, average prices, total sales, industry profits and consumer surplus are (almost) identical under monopoly and duopoly.[51] This result is similar to Meisner (2017), who shows that under similar conditions but with different consumer heterogeneity the competitive price path resembles the monopoly price path if market capacities are equal and scarce.

[50]To further strengthen this result, note that my forward-looking consumer model is rather 'conservative' in the sense that only one consumer can buy in a period, imposing a substantial rationing risk.

[51]Note that if capacity is very scarce, e.g. if $x_1^D = x_2^D = x^M/2 = 4$ in $t = 10$, total expected sales might even be slightly lower under duopoly than under duopoly as pricing under duopoly can be less aggressive than under monopoly where the positive externality on other capacities is internalized (cf. Section 5.6).

Figure 18: Competition Policy



Parameters: $t = 10$, $v = 100$, $\delta = 1$, $\mu = 10$, myopic consumers. Average prices (a), total sales (b), industry profits (c) and consumer surplus (d), as expected in $t = 10$, under monopoly and (initially symmetric) duopoly for different levels of initial market capacity $\sum_i x_i$, such that $x_1^D = x_2^D = x^M/2$.

What is more, in Figure 18 I show that if total market capacity becomes excessive relative to remaining time, i.e. $\sum_i x_i > t$, average prices under duopoly are lower than under monopoly. Consequently, under duopoly there will be more total sales until the deadline, while consumer surplus becomes larger and industry profits smaller than under monopoly. Further, under monopoly for $x^M = \sum_i x_i > t$ and respectively under duopoly for $x_i^D = \sum_i x_i/2 > t$, all four measures become constant in additional market capacity $\sum_i x_i$ because a situation of individual excess capacity is reached.[52] Summing up, we get

[52]For duopoly asymmetry, refer to Figure A.8 of Appendix A.3, where I show that for extremely as well as for less extremely asymmetric initial capacities all results are robust.

the following result.

Result 14. *If market capacity is excessive, average prices and industry profits are typically higher under monopoly than under duopoly, while consumer surplus and total sales are typically lower. If market capacity is scarce, duopoly and monopoly results are relatively equal.*

This result is robust with forward-looking consumers. To see this, consider Figure A.9 in Appendix A.3 for average prices and total sales with forward-looking consumers.

6.4 Ex-ante Capacity Production

In the previous sections of this paper all analyses were conducted for exogenously given levels of capacity. In this section I will consider endogenous capacity production before the actual dynamic pricing game starts. I will compare equilibrium capacity production with efficient capacity production as well as capacity choices under capacity production collusion and under monopoly.

Suppose that before they start selling, firms can build capacities $\hat{x}_i$ at some (constant) unit cost $k \cdot v$, whereby $k \leq 1$ shall represent the fraction of the good's value v. To evaluate firms' capacity production for different levels of production costs k, let firm profits in state $\boldsymbol{\omega}$ be as defined in Section 6.1, minus capacity costs per unit of production, i.e. $\Pi_i(\boldsymbol{\omega}) - \hat{x}_i\, k\, v$.[53] Then joint profits are the sum over all firms' profits, i.e. $\sum_i \{\Pi_i(\boldsymbol{\omega}) - \hat{x}_i\, k\, v\}$. Moreover, I define collusion in the sense that firms only collude in capacity production, i.e. produce capacities which jointly maximize industry profits in the (expected) subsequent dynamic pricing game minus initial capacity production costs, while pricing remains non-cooperative as before. Further, to define an efficiency benchmark with capacity costs, consider the number of total sales in a market, as given in Section 6.1. Capacity costs reduce the value of all produced capacities by the factor of k, hence the efficiency measure in state $\boldsymbol{\omega}$ will be defined by $S(\boldsymbol{\omega}) - \sum_i \hat{x}_i\, k$. Then capacities are efficient if they maximize total sales in the subsequent dynamic pricing game, where firms are free to post prices, minus the initial capacity value costs.

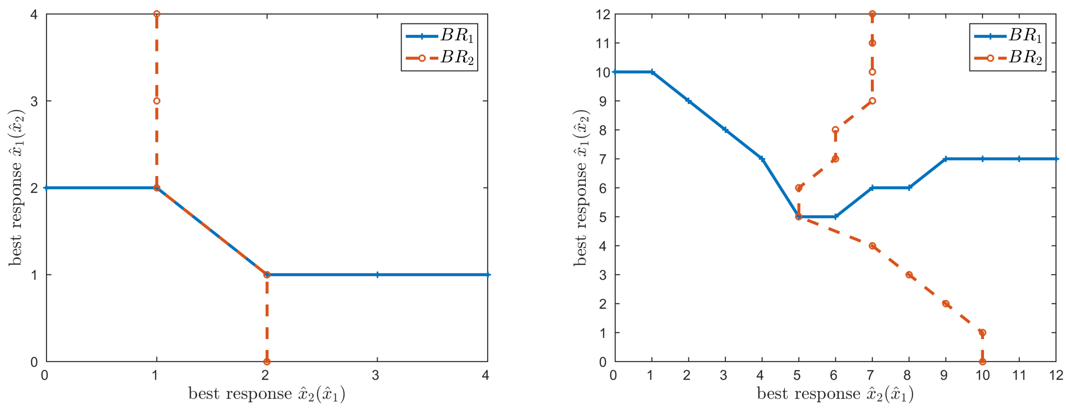
When producing capacities, firms consider their expected profits given their own and their competitiors' capacity choices. Thus, in the capacity production equilibrium firms' capacity choices are mutual best responses. Generally, multiple and mixed-strategy equilibria are possible. In Figure 19 the capacity production best responses are given for

[53] On average, even under monopoly, firms will sell at a price below v, hence subtracting a fraction k of v exhibits a sensible production cost assumption.

subsequent pricing games of length $T = 2$ and $T = 10$, while capacity production costs are $k = 0$. For $T = 2$ equilibria consist of $\hat{x}_i = 1$ and $\hat{x}_{-i} = 2$.[54] Contrary, for $T = 10$ the equilibrium is uniquely given by the pure strategies $\hat{x}_i = \hat{x}_{-i} = T/2$.

FIGURE 19: Capacity Production: Best Responses and Equilibrium

(a) $T = 2$ (b) $T = 10$



Parameters: $v = 100$, $\delta = 1$, $\mu = 10$, $k = 0$, myopic consumers, duopoly. Best responses in ex-ante capacity production. In (a) $T = 2$, in (b) $T = 10$.

As long as competitor capacity $\hat{x}_{-i}$ is small relative to remaining time, firm i will (weakly) decrease its capacity $\hat{x}_i$ in $\hat{x}_{-i}$, i.e. capacities are strategic substitutes. For higher levels of $\hat{x}$ firm i might even expand $\hat{x}_i$ in $\hat{x}_{-i}$, s.t. capacities are strategic complements, but only to an extent such that i remains the smaller firm in the new situation of aggregate excess capacities.[55] In equilibrium, firms will choose to produce sufficient capacities to jointly cover the entire expected demand, but without obtaining (too much) aggregate excess capacity, which would trigger equilibrium paths of consistently lower prices.[56] Note from Dudey (1992) that for $\mu \to 0$ capacity choices in duopoly are always such that exactly $\sum_i \hat{x}_i = T$.[57] Summing up, the introduction of consumer heterogeneity $\mu > 0$ can yield (small) excess capacity production.

[54] In $t = 2$ firm $-i$ prefers to build one excess capacity because then both firms' prices will be more aggressive such that the probability of a sale by firm i increases and firm $-i$ will remain with a higher probability as the monopolist in $t = 1$.

[55] Note that also in Dudey's (1992) model of dynamic Bertrand an entrant would like to just undercut the incumbent's capacity choice, if there are aggregate excess capacities.

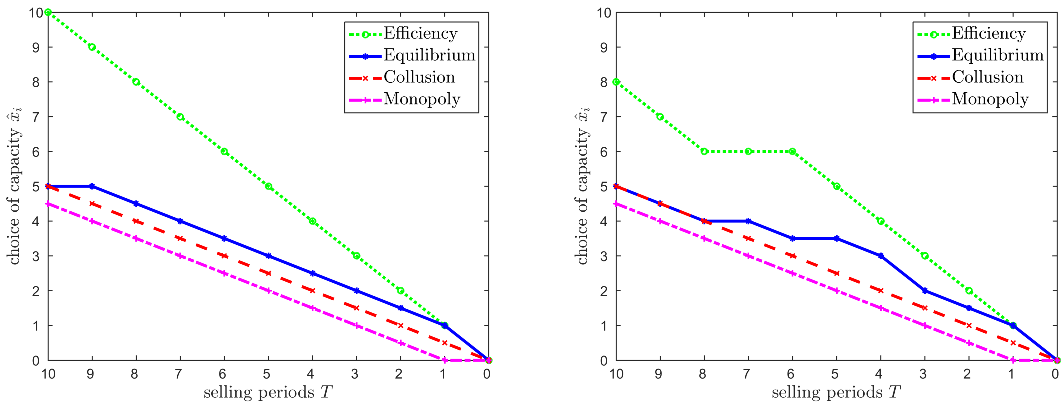
[56] If T is very high or μ relatively high, the probability that in some periods no trade happens increases, such that firms might produce less capacities than remaining time periods.

[57] There, if t is even, $\hat{x}_1 = \hat{x}_2 = t/2$ and if t is odd, $\hat{x}_i = (T + 1)/2$ while $\hat{x}_{-i} = (T - 1)/2$.

Result 15. *In the ex-ante equilibrium of capacity production without costs firms might build excess capacities in equilibrium.*

FIGURE 20: Comparison of Capacity Choices without Costs

(a) Myopic Consumers (b) Forward-looking Consumers



Parameters: $v = 100$, $\delta = 1$, $\mu = 10$, $k = 0$ duopoly for (a) myopic consumers and (b) forward-looking consumers with $\delta_c = \delta = 1$. Capacity choices for a firm $\hat{x}_i$ under efficiency, in equilibrium, under collusion and under monopoly for all selling periods T. For multiple equilibria $\hat{x}_i$ is the mean of their pure-strategies. Under monopoly I plot $\hat{x}_i = \hat{x}_M/2$ to compare.

Consider Figure 20 to see capacity choices $\hat{x}_i$ under duopoly with zero capacity costs in the four different regimes, i.e. under efficiency, in equilibrium, under collusion and under monopoly, for all possible subsequent selling horizons $T \leq 10$, and for myopic (a) as well as forward-looking (b) consumers. Note for all multiple equilibria that the sum of both firms' capacities is still unique in all cases I considered, such that I take the mean capacity of a pure-strategy equilibrium.[58] Under monopoly and collusion firms restrict themselves to $\hat{x}_M \leq T$ and $\hat{x}_i \leq T/2$ respectively, while capacities in equilibrium are such that $\hat{x}_i \geq T/2$ because the (negative) externality on the other firm's profit is not internalized.[59] Contrary, it would be efficient to have individually excess capacities, i.e. $\hat{x}_i \geq T$ for all i, to ensure lowest prices and hence highest total sales. Interestingly, when consumers are forward-looking more or less capacity than with myopic consumers could be built in equilibrium. This is because with forward-looking consumers on the one hand due to the waiting option less demand is expected but on the other hand more capacity and hence lower prices reduce the relative value of waiting. Nevertheless, the order of capacity choices remains robust.

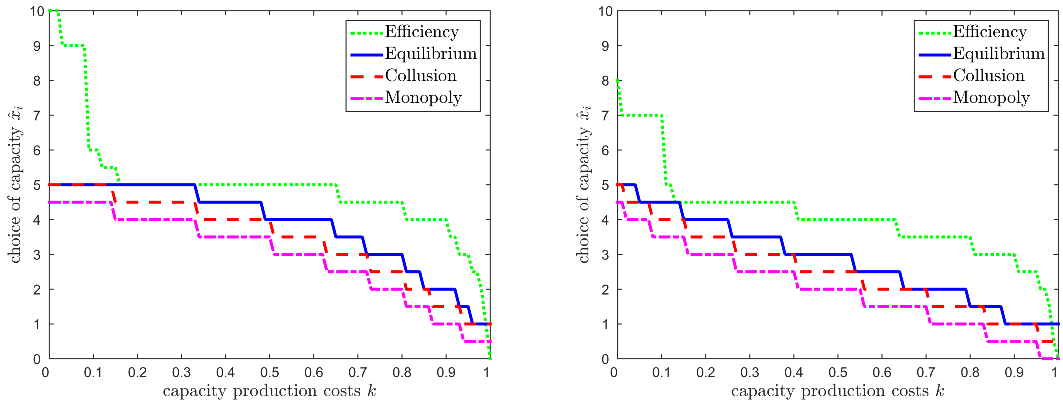
[58] Thereby I neglect the mixed-strategy equilibrium, which however yields the same aggregate capacities.
[59] Note footnote 56 to see why $\hat{x}_i < T/2$ is possible for high T or high μ.

Result 16. *In the ex-ante equilibrium of capacity production capacities are smaller than or equal to the efficient level, yet at least as large as under collusion or even monopoly.*

To see that this is also robust to different capacity costs, consider Figure 21, which shows all four capacity choice regimes for different levels of k and $T = 10$.[60]

Figure 21: Comparison of Capacity Choices for Different Costs

(a) Myopic Consumers (b) Forward-looking Consumers



Parameters: $T = 10$, $v = 100$, $\delta = 1$, $\mu = 10$, duopoly for (a) myopic consumers and (b) forward-looking consumers with $\delta_c = \delta = 1$. Capacity choices for a firm $\hat{x}_i$ under efficiency, in equilibrium, under collusion and under monopoly for different capacity production costs k. For multiple equilibria $\hat{x}_i$ is the mean of their pure-strategies. Under monopoly I plot $\hat{x}_i = \hat{x}_M/2$ to compare.

Finally, the following observations further strengthen my results on policies allowing consumers to become forward-looking. For this I consider the ex-ante capacity production equilibrium with myopic and forward-looking consumers and then compare welfare measures of the subsequent dynamic pricing games under both regimes.[61] Consider Figure 22, which shows (a) efficiency (total sales) and (b) average prices for $T = 10$ under both regimes for different production costs k.[62] Efficiency and average prices are higher with capacity equilibrium choices and myopic consumers, if capacity costs are not too large. Compare Figure A.10 from Appendix A.3 to see that consequently also industry

[60]For extremely large capacity costs it would be efficient not to produce any capacity, however firms might still expect some positive valuation shock during any of the remaining time periods, such that they might sell at a price above v and hence find it profitable to build one capacity.

[61]Note that here for all multiple equilibria the sum of capacities is still unique such that I can use the results of any of the pure-strategy equilibria as they are symmetric.

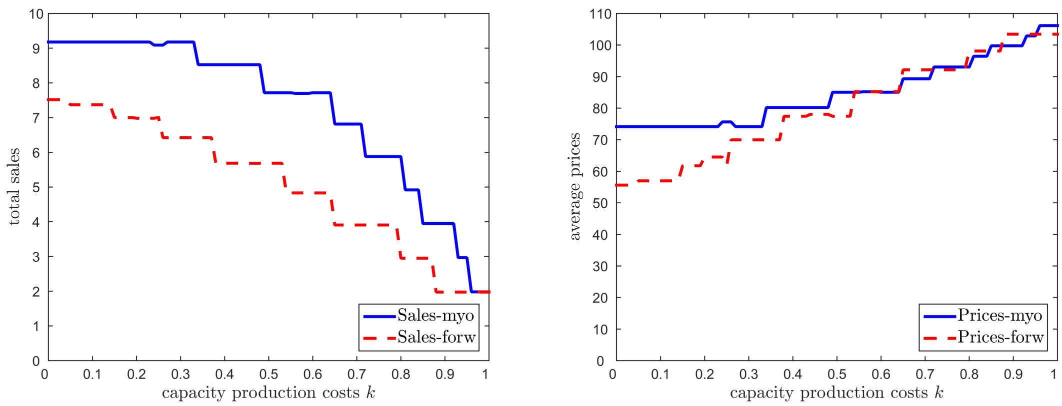
[62]Note that for Figures 22 and A.10 I do not subtract the initial capacity costs for reasons of comparability, however the results would remain qualitatively robust. It should be noted though, that the equilibrium capacity production choices could represent a Prisoners' Dilemma and lead to an efficiency loss and also to negative industry profits for high capacity costs k.

profits are higher while consumer surplus is lower, as long as capacity costs are not too large.[63] From this we can conclude that if we consider the ex-ante game of capacity production, the results from the comparison of forward-looking and myopic consumers from Section 6.2 carry over, if capacity costs are not too large.

Result 17. *Given ex-ante equilibrium capacity production choices with myopic and forward-looking consumers, total sales, average prices and industry profits are higher with myopic consumers, while consumer surplus is lower than with forward-looking consumers, as long as capacity costs are not too large.*

FIGURE 22: Welfare Measures Given the Ex-ante Capacity Equilibrium

(a) Total Sales (b) Average Prices



Parameters: $T = 10$, $v = 100$, $\delta = 1$, $\mu = 10$, duopoly for myopic consumers and forward-looking consumers with $\delta_c = \delta = 1$. Total sales (a) and average price (b) of the dynamic pricing game after ex-ante equilibrium capacity production, for different k. For multiple equilibria I consider the mean of their pure-strategies.

[63] Note footnote 62. Further note that while in $T = 10$ equal capacities are built in equilibrium with myopic and with forward-looking consumers for $k = 0$, this result is also robust for T where capacity production might differ in both cases. Consider Figure A.11 in Appendix A.3 and note that e.g. in $T = 5$, although more capacity might be produced with forward-looking consumers, total sales are still lower. Similarly, although in $T = 8$ more capacity might be produced under consumer myopia, consumers are still better off when forward-looking, as long as k is not too large.

7 Conclusion

I study dynamic pricing of capacity-constrained firms under oligopolistic competition. For this I employ a multi-period model with heterogeneous consumer demand such that firms have an inter-temporal pricing problem, while forward-looking consumers face an inter-temporal waiting problem. The resulting pure-strategy price path equilibrium can explain empirically observed price volatility, which does not arise in other models of the literature. Crucially, the number of remaining capacities of each oligopolist as well as the total number of market capacities relative to remaining selling time determine firms' reservation values and hence their prices in a given period. I find that price-leadership is not monotone in capacity-leadership and competitive prices are also neither monotone in a firm's capacity nor in remaining time, unless capacities are scarce.

There are three main policy results for these dynamic pricing markets. First, a policy allowing consumers to become forward-looking increases consumer surplus but decrease efficiency (total sales) as well as industry profits because of the increased pressure on prices. Second, stronger competition policy is especially valuable if market capacities are excessive relative to expected demand. And third, ex-ante equilibrium capacity production can be excessive, though still inefficiently small, while under capacity production collusion or under monopoly even less capacity is built.

Building upon this study, further work could refine on a consumer equilibrium, e.g. allow for more than one consumer to strategically compete for sales in a period. Additionally, waiting consumers could have persisting taste shocks. Further, it would be interesting to investigate the intermediate case of forward-looking consumers who do not have knowledge about firms' capacity levels. Also, collusion through repeated interaction in capacity production and dynamic pricing could be studied. Additionally, the analysis of price commitment and fixed prices in this setting could be worthwhile investigating for many real-world dynamic pricing applications. Finally, many dynamics are in effect at the same time and this model presents an approach to disentangle them. It could be a valuable empirical exercise to test these predictions with, e.g., real-world airfare data.

A Appendix

A.1 Myopic Consumers - Baseline Model

Demand for $\mu \to 0$

$$D_i(\mathbf{p},\boldsymbol{\omega}) = \frac{1}{\frac{1}{\exp\left(\frac{v-p_i}{\mu}\right)} + \frac{\exp\left(\frac{v-p_i}{\mu}\right)}{\exp\left(\frac{v-p_i}{\mu}\right)} + \frac{\sum_{j\neq i}\exp\left(\frac{v-p_j}{\mu}\right)}{\exp\left(\frac{v-p_i}{\mu}\right)}} = \frac{1}{\frac{1}{\exp\left(\frac{v-p_i}{\mu}\right)} + 1 + \sum_{j\neq i}\exp\left(\frac{p_i-p_j}{\mu}\right)}$$

First, note that if $p_i > v$, the whole expression tends to zero for $\mu \to 0$ because the first term in the denominator tends to ∞. Second, let $p_i < v$.[64] Then for $\mu \to 0$ the first term in the denominator tends to zero. If $p_i > p_j$ for at least one $j \neq i$, then for $\mu \to 0$ at least one of the summands in the third term of the denominator tends to ∞ and hence the whole demand expression to 0. If $p_i < p_j$ for all $j \neq i$, then for $\mu \to 0$ the third term (summation) tends to $(n-1)0$ and hence the whole expression to 1. If $p_i = p_j$ for all $j \neq i$, then for $\mu \to 0$ the third third term (summation) equals $(n-1)1$ and the whole demand to $1/n$. Hence,

$$\lim_{\mu\to 0} D_i(\mathbf{p},\boldsymbol{\omega}) = \begin{cases} 0 & \text{if } p_i > v \\ 0 & \text{if } p_i < v \text{ and } p_i > p_j \text{ for at least one } j \neq i \\ 1 & \text{if } p_i < v \text{ and } p_i < p_j \text{ for all } j \neq i \\ 1/n & \text{if } p_i < v \text{ and } p_i = p_j \text{ for all } j \neq i. \end{cases}$$

Derivatives of Demand

For all $i, j \in J(\boldsymbol{\omega})$ and $j \neq i$ we have

$$\frac{\partial D_0(\mathbf{p},\boldsymbol{\omega})}{\partial p_i} = \frac{0 - [1]\left[\exp\left(\frac{v-p_i}{\mu}\right)\left(-\frac{1}{\mu}\right)\right]}{\left[1 + \sum_{J(\boldsymbol{\omega})}\exp\left(\frac{v-p_j}{\mu}\right)\right]^2} = \frac{1}{\mu}D_i(\mathbf{p},\boldsymbol{\omega})D_0(\mathbf{p},\boldsymbol{\omega}) > 0.$$

[64] If $p_i = v < p_j$, for all $j \neq i$, demand is equal to $1/2$ and if $p_i = p_j = v$, demand is equal to $1/(n+1)$, because the outside option utility is normalized to zero, too.

$$
\begin{aligned}
\frac{\partial D_i(\mathbf{p},\boldsymbol{\omega})}{\partial p_i} &= D_i(\mathbf{p},\boldsymbol{\omega}) \frac{\left[1+\sum_{J(\boldsymbol{\omega})} \exp\left(\frac{v-p_j}{\mu}\right)\right]\left(-\frac{1}{\mu}\right) - \left[\exp\left(\frac{v-p_i}{\mu}\right)\left(-\frac{1}{\mu}\right)\right]}{1+\sum_{J(\boldsymbol{\omega})} \exp\left(\frac{v-p_j}{\mu}\right)} \\
&= -\frac{1}{\mu} D_i(\mathbf{p},\boldsymbol{\omega}) \frac{1+\sum_{J(\boldsymbol{\omega})\setminus\{i\}} \exp\left(\frac{v-p_j}{\mu}\right)}{1+\sum_{J(\boldsymbol{\omega})} \exp\left(\frac{v-p_j}{\mu}\right)} = -\frac{1}{\mu} D_i(\mathbf{p},\boldsymbol{\omega})\left[1-D_i(\mathbf{p},\boldsymbol{\omega})\right] < 0. \\
\frac{\partial D_i(\mathbf{p},\boldsymbol{\omega})}{\partial \mu} &= \frac{\exp\left(\frac{v-p_i}{\mu}\right)\left\{\left(-\frac{v-p_i}{\mu^2}\right) + \sum_{J(\boldsymbol{\omega})}\left[\exp\left(\frac{v-p_j}{\mu}\right)\left(\frac{p_i-p_j}{\mu^2}\right)\right]\right\}}{\left[1+\sum_{J(\boldsymbol{\omega})} \exp\left(\frac{v-p_j}{\mu}\right)\right]^2} \\
&\leq 0 \text{ if } p_i < v \text{ and } p_i < p_j \text{ for all } j \neq i. \\
\frac{\partial D_0(\mathbf{p},\boldsymbol{\omega})}{\partial \mu} &= \frac{0-[1]\sum_{J(\boldsymbol{\omega})}\left[\exp\left(\frac{v-p_j}{\mu}\right)\left(-\frac{v-p_j}{\mu^2}\right)\right]}{\left[1+\sum_{J(\boldsymbol{\omega})} \exp\left(\frac{v-p_j}{\mu}\right)\right]^2} \geq 0 \text{ if } p_j < v \text{ for all } j \neq i.
\end{aligned}
$$

Derivatives of Value Function $V_i(\mathbf{p},\boldsymbol{\omega}) = D_i(\mathbf{p},\boldsymbol{\omega})\, p_i + \delta \sum_j \left\{D_j(\mathbf{p},\boldsymbol{\omega})\, W_{i,j}(\boldsymbol{\omega})\right\}$

Using simpler notation and $j \in J(\boldsymbol{\omega}) \cup \{0\}$ we get the first derivatives

$$
\begin{aligned}
\frac{\partial V_i(\mathbf{p},\boldsymbol{\omega})}{\partial p_i} &= \left(-\frac{1}{\mu}\right) D_i\left[1-D_i\right] p_i + D_i + \frac{\delta}{\mu} D_i \left[\sum_{j\neq i}\left\{D_j W_{i,j}\right\} - (1-D_i) W_{i,i}\right] \\
&= D_i\left[1-\frac{1}{\mu}p_i - \frac{\delta}{\mu}W_{i,i} + \frac{1}{\mu}V_i(\mathbf{p},\omega)\right] \\
&= D_i\left[1-\frac{1}{\mu}\sum_{j\neq i} D_j\left[p_i+\delta W_{i,i}-\delta W_{i,j}\right]\right]. \\
\frac{\partial V_i(\mathbf{p},\boldsymbol{\omega})}{\partial p_j} &= \frac{1}{\mu} D_j\left[D_i p_i + \delta \sum_{k\neq j}\left\{D_k W_{i,k}\right\} - (1-D_j)\delta W_{i,j}\right] \\
&= \frac{1}{\mu} D_j\left[p_i + \delta W_{i,i} - \delta W_{i,j} - \sum_{j\neq i} D_j\left(p_i + \delta\, W_{i,i} - \delta\, W_{i,j}\right)\right] \\
&= \frac{1}{\mu} D_j\left[V_i - \delta W_{i,j}\right].
\end{aligned}
\tag{A.1}
$$

Second Derivatives of Value Function

$$\begin{aligned}\frac{\partial^2 V_i(\mathbf{p},\boldsymbol{\omega})}{\partial p_i^2} &= \frac{\partial D_i}{\partial p_i}\left[1-\frac{1}{\mu}p_i-\frac{\delta}{\mu}W_{i,i}+\frac{1}{\mu}V_i(\mathbf{p},\boldsymbol{\omega})\right]+D_i\left[-\frac{1}{\mu}+\frac{1}{\mu}\frac{\partial V_i(\mathbf{p},\boldsymbol{\omega})}{\partial p_i}\right]\\ &= \frac{\partial V_i(\mathbf{p},\boldsymbol{\omega})}{\partial p_i}\left[2D_i-1\right]-\frac{1}{\mu}D_i.\end{aligned}$$

$$\begin{aligned}\frac{\partial^2 V_j(\mathbf{p},\boldsymbol{\omega})}{\partial p_j \partial p_i} &= \frac{\partial D_j}{\partial p_i}\left[1-\frac{1}{\mu}p_j-\frac{\delta}{\mu}W_{j,j}+\frac{1}{\mu}V_j(\mathbf{p},\boldsymbol{\omega})\right]+D_j\left[\frac{1}{\mu}\frac{\partial V_j(\mathbf{p},\boldsymbol{\omega})}{\partial p_i}\right]\\ &= \frac{1}{\mu}D_i\frac{\partial V_j(\mathbf{p},\boldsymbol{\omega})}{\partial p_j}+\frac{1}{\mu^2}D_jD_i\left[V_j(\mathbf{p},\boldsymbol{\omega})-\delta W_{j,i}(\boldsymbol{\omega})\right].\end{aligned}$$

Proof of Lemma 1[65]

Proof. Define $w_{ij}(\boldsymbol{\omega}) \equiv W_{i,j}(\boldsymbol{\omega})-W_{i,i}(\boldsymbol{\omega})$, $w_i^+(\boldsymbol{\omega}) \equiv \max_j w_{ij}(\boldsymbol{\omega})$ and $w_i^-(\boldsymbol{\omega}) \equiv \min_j w_{ij}(\boldsymbol{\omega})$, where $i,j \in J(\boldsymbol{\omega}) \cup \{0\}$ and $i \neq j$. Note that $w_i^+(\boldsymbol{\omega})$ and $w_i^-(\boldsymbol{\omega})$ are constant in a given state $\boldsymbol{\omega}$.[66] Then, from (10) we get that

$$\frac{\partial V_i(\mathbf{p},\boldsymbol{\omega})}{\partial p_i} > D_i(\mathbf{p},\boldsymbol{\omega})\left\{1-\frac{1}{\mu}\Big(1-D_i(\mathbf{p},\boldsymbol{\omega})\Big)\Big[p_i-\delta w_i^-(\boldsymbol{\omega})\Big]\right\}. \tag{A.2}$$

Note for the lower bound that the right-hand-side of (A.2) is positive, if $p_i < \frac{\mu}{1-D_i(\mathbf{p},\boldsymbol{\omega})} + \delta w_i^-(\boldsymbol{\omega})$. Since $\mu + \delta w_i^-(\boldsymbol{\omega}) \leq \frac{\mu}{1-D_i(\mathbf{p},\boldsymbol{\omega})} + \delta w_i^-(\boldsymbol{\omega})$, this is in particular satisfied, if

$$p_i < \mu + \delta w_i^-(\boldsymbol{\omega}). \tag{A.3}$$

For the upper bound, we get from equation (10) that

$$\frac{\partial V_i(\mathbf{p},\boldsymbol{\omega})}{\partial p_i} < D_i(\mathbf{p},\boldsymbol{\omega})\left\{1-\frac{1}{\mu}\Big(1-D_i(\mathbf{p},\boldsymbol{\omega})\Big)\Big[p_i-\delta w_i^+(\boldsymbol{\omega})\Big]\right\}. \tag{A.4}$$

Note that the right-hand-side of (A.4) is negative, if $p_i > \frac{\mu}{1-D_i(\mathbf{p},\boldsymbol{\omega})} + \delta w_i^+(\boldsymbol{\omega})$. Since

$$1-D_i(\mathbf{p},\boldsymbol{\omega}) = 1-\frac{\exp(\frac{v-p_i}{\mu})}{1+\sum_{J(\boldsymbol{\omega})}\exp(\frac{v-p_j}{\mu})} \geq 1-\frac{\exp(\frac{v-p_i}{\mu})}{1+\exp(\frac{v-p_i}{\mu})} \geq 1-\frac{\exp(\frac{v}{\mu})}{1+\exp(\frac{v}{\mu})} = \frac{1}{1+\exp(\frac{v}{\mu})}$$

[65]For parts of Lemma 1 I adopt the construction of a similar proof in Lin and Sibdari (2009).

[66]This is true because continuation equilibrium multiplicity is excluded, as seen in the proof of Proposition 1.

we have that $\mu\left(1+\exp(\frac{v}{\mu})\right)+\delta w_i^+(\boldsymbol{\omega}) \geq \frac{\mu}{1-D_i(\mathbf{p},\boldsymbol{\omega})}+\delta w_i^+(\boldsymbol{\omega})$. Therefore the right-hand-side of (A.4) is in particular negative, if

$$p_i > \mu\left(1+\exp(\frac{v}{\mu})\right)+\delta w_i^+(\boldsymbol{\omega}). \tag{A.5}$$

Hence $V_i(\mathbf{p},\omega)$ increases in p_i for $p_i < \mu+\delta w_i^-(\boldsymbol{\omega})$ and decreases in p_i for $p_i > \mu(1+\exp(\frac{v}{\mu}))+\delta w_i^+(\boldsymbol{\omega})$, given any prices $\mathbf{p}_{-i}$ of the competitors. Thus, no other prices need to be considered and the best-response price is bounded, i.e. inf $p_i(\mathbf{p_{-i}},\boldsymbol{\omega}) \geq \mu+\delta w_i^-(\boldsymbol{\omega}) > -\infty$ and sup $p_i(\mathbf{p_{-i}},\boldsymbol{\omega}) \leq \mu(1+\exp(\frac{v}{\mu}))+\delta w_i^+(\boldsymbol{\omega}) < \infty$. □

Proof of Lemma 2

Proof. Whenever the first-order condition given by setting expression (9) equal to zero holds, i.e. whenever $\partial V_i(\mathbf{p},\boldsymbol{\omega})/\partial p_i = 0$, it follows for equation (11) that

$$\frac{\partial^2 V_i(\mathbf{p},\boldsymbol{\omega})}{\partial p_i^2} = -\frac{1}{\mu}D_i\left(\mathbf{p},\boldsymbol{\omega}\right) < 0, \tag{A.6}$$

i.e. the second-order condition for a local maximum is fulfilled. Note that the strict inequality follows from the fact that $D_i\left(\mathbf{p},\boldsymbol{\omega}\right) > 0$ because prices are bounded from above, which we showed in Lemma 1. This satisfies the sufficient condition for quasi-concavity as stated in Crouzeix (1980), which is an extension of Katzner's criterion (Katzner et al., 1970). □

Proof of Proposition 1

Proof. Following Fudenberg and Tirole (1991), who rely on e.g. Debreu (1952), a pure-strategy Nash equilibrium exists in each period t if the strategy sets are (a) non-empty, (b) compact and (c) convex, while the payoff to a firm is (d) continuous in all firms' actions and (e) quasi-concave in its own action.

(a) and (c) are fulfilled since firms choose from the non-empty and convex set $p_i \in \mathbb{R}$. From Lemma 1 we know that there is an upper and a lower limit for firms' best response pricing functions. Hence we could construct a new but equivalent game where firms are restricted to prices $p_i \in [\mu+\delta w_i^-(\boldsymbol{\omega}),\ \mu\left(1+\exp(\frac{v}{\mu})\right)+\delta w_i^+(\boldsymbol{\omega})]$, such that we also have (b) a compact strategy set. To see that a firm's payoff function $V_i(\mathbf{p},\boldsymbol{\omega})$ is continuous in all firms' actions (d), note that equation (6) is continuous in all p_j, $j \in \{1,...,n\}$, since $W_{i,j}(\boldsymbol{\omega})$ and $W_{i,0}(\boldsymbol{\omega})$ are constant in $\mathbf{p}$.[67] Finally, from Lemma 2 we get (e), payoff quasi-concavity in p_i. Therefore the Nash equilibrium is in pure strategies as the set of strategies consists only of pure strategies.

If there were multiple such equilibria in a given period, then I assume that firms will coordinate on the equilibrium granting higher total payoffs.[68]

Under monopoly only one firm, say i, holds capacity. Then from the first-order condition $\partial V_i(\mathbf{p}, \boldsymbol{\omega})/\partial p_i = 0$, it follows that $\partial^2 V_i(\mathbf{p}, \boldsymbol{\omega})/\partial p_i^2 = -\frac{1}{\mu} D_i(\mathbf{p}, \boldsymbol{\omega}) < 0$, and hence the monopoly price choice exists and is unique.

Hence, there exists a pure-strategy Nash equilibrium in each period and firms can perfectly anticipate future equilibrium outcomes, while in the final period $t = 1$ there are no continuation values. Then, by backward induction there exists a pure-strategy Sub-game Perfect Nash Equilibrium for the whole game. □

Proof of Condition 1

Proof. To show uniqueness I use the index theory approach (Vives, 2001) based on the Poincaré-Hopf index theorem. For this define the marginal value function as $g_i(\mathbf{p}) \equiv \partial V_i(\mathbf{p}, \boldsymbol{\omega})/\partial p_i$ for all firms $i \in \{1, 2\}$. Let $g(\mathbf{p}) = (g_1(\mathbf{p}), g_2(\mathbf{p}))$. Define the Jacobian matrix of $g(\mathbf{p})$ as $M[g(\mathbf{p})]$. The above named theorem implies that if $g : P \to \mathbb{R}$, where P is a compact set in $\mathbb{R}$, satisfies a boundary condition and the determinant of $M[-g(\mathbf{p})]$ is positive whenever $g(\mathbf{p}) = 0$, then there is a unique solution to $g(\mathbf{p}) = 0$.

From Lemma 1 we know that the best-response prices are bounded in $[b_i^-, b_i^+] := [\mu + \delta w_i^-(\boldsymbol{\omega}), \mu \left(1 + \exp(\frac{v}{\mu})\right) + \delta w_i^+(\boldsymbol{\omega})]$ for each i, hence we have a compact set. For the boundary condition note that $\partial V_i(b_i^-, p_j, \boldsymbol{\omega})/\partial p_i > 0$ and $\partial V_i(b_i^+, p_j, \boldsymbol{\omega})/\partial p_i < 0$ for all p_j and all $i, j \in \{1, 2\}$, $j \neq i$, which is shown in the proof of Lemma 1. Then the Jacobian matrix of $-g(\mathbf{p})$ is given by

$$M[-g(\mathbf{p})] = \begin{pmatrix} -\partial^2 V_1(\mathbf{p}, \omega)/\partial p_1^2 & -\partial^2 V_1(\mathbf{p}, \omega)/\partial p_1 \partial p_2 \\ -\partial^2 V_2(\mathbf{p}, \omega)/\partial p_2 \partial p_1 & -\partial^2 V_2(\mathbf{p}, \omega)/\partial p_2^2 \end{pmatrix}.$$

The determinant of this matrix, given the above specified second derivatives, is
$\det M[-g(\mathbf{p})]$

[67]This is true because continuation equilibrium multiplicity is excluded, as seen below.

[68]Refer to the discussion of equilibrium uniqueness in the Section 3.3 but note that during simulations for no parameter constellation equilibrium multiplicity was encountered.

$$
\begin{aligned}
&= \left[\frac{\partial^2 V_1(\mathbf{p},\boldsymbol{\omega})}{\partial p_1^2}\frac{\partial^2 V_2(\mathbf{p},\boldsymbol{\omega})}{\partial p_2^2}\right] - \left[\frac{\partial^2 V_1(\mathbf{p},\boldsymbol{\omega})}{\partial p_1 \partial p_2}\frac{\partial^2 V_2(\mathbf{p},\boldsymbol{\omega})}{\partial p_2 \partial p_1}\right] \\
&= \left[\frac{1}{\mu}\frac{\partial V_1(.)}{\partial p_1}\left[2D_1(.)-1\right] - \frac{1}{\mu}D_1(.)\right]\left[\frac{1}{\mu}\frac{\partial V_2(.)}{\partial p_2}\left[2D_2(.)-1\right] - \frac{1}{\mu}D_2(.)\right] \\
&\quad - \left[\frac{1}{\mu}D_2(.)\frac{\partial V_1(.)}{\partial p_1} + \frac{1}{\mu}D_1(.)\frac{\partial V_1(.)}{\partial p_2}\right]\left[\frac{1}{\mu}D_1(.)\frac{\partial V_2(.)}{\partial p_2} + \frac{1}{\mu}D_2(.)\frac{\partial V_2(.)}{\partial p_1}\right].
\end{aligned}
$$

Whenever $g(\mathbf{p}) = 0$, i.e. $\partial V_i(\mathbf{p},\boldsymbol{\omega})/\partial p_i = 0\ \forall i$, we get that

$$
M\left[-g(\mathbf{p})\right] = \frac{1}{\mu^2}D_1(\mathbf{p},\boldsymbol{\omega})\,D_2(\mathbf{p},\boldsymbol{\omega})\left[1 - \frac{\partial V_1(\mathbf{p},\boldsymbol{\omega})}{\partial p_2}\frac{\partial V_2(\mathbf{p},\boldsymbol{\omega})}{\partial p_1}\right].
$$

Thus, if $\partial V_i(\mathbf{p},\omega)/\partial p_j \in (-1,\,1)$ while $\partial V_i(\mathbf{p},\omega)/\partial p_i = 0$, we have $M\left[-g(\mathbf{p})\right] > 0$ and we can apply the Poincaré-Hopf index theorem, hence a unique solution to the first-order conditions exists. Finally, note from equation (A.1) that with (12) and (13) we get

$$
\left.\frac{\partial V_i(\mathbf{p},\boldsymbol{\omega})}{\partial p_j}\right|_{\frac{\partial V_i(\mathbf{p},\omega)}{\partial p_i}=0} = \frac{1}{\mu}D_j(\mathbf{p},\boldsymbol{\omega})\left[p_i + \delta W_{i,i}(\boldsymbol{\omega}) - \delta W_{i,j}(\boldsymbol{\omega}) - \mu\right]. \qquad \text{(A.7)}
$$

□

Proof of Proposition 2

Proof. For uniqueness we need Condition 1 to hold. From equation (A.7) together with (12) we get that

$$
\frac{\partial V_i(\mathbf{p},\boldsymbol{\omega})}{\partial p_j}\Big|_{\frac{\partial V_i(\mathbf{p},\omega)}{\partial p_i}=0} = 1 - D_j(\mathbf{p}^*,\boldsymbol{\omega}) - \frac{1}{\mu}D_0(\mathbf{p}^*,\boldsymbol{\omega})\left[p_i^* + \delta W_{i,i}(\boldsymbol{\omega}) - \delta W_{i,0}(\boldsymbol{\omega})\right] \quad \text{(A.8)}
$$

Note for $\mu \to 0$, using L'Hôpital's Rule, that

$$
\lim_{\mu\to 0}\frac{1}{\mu}D_0(\mathbf{p},\boldsymbol{\omega}) = \lim_{\mu\to 0}\frac{1/\mu}{1 + \sum\limits_{J(\boldsymbol{\omega})}\exp\left(\frac{v-p_j}{\mu}\right)} = \lim_{\mu\to 0}\frac{1}{\sum\limits_{J(\boldsymbol{\omega})}\left\{(v-p_j)\exp\left(\frac{v-p_j}{\mu}\right)\right\}} = 0.
$$

Consequently, for $\mu \to 0$ the third term in equation (A.8) tends to zero since p_i^*, $W_{i,i}(\boldsymbol{\omega})$ and $W_{i,0}(\boldsymbol{\omega})$ are then bounded by v and $t \cdot v$ respectively, as a firm i only obtains positive demand if $p_i \le v$. Then, since $1 - D_j(\mathbf{p},\boldsymbol{\omega}) \in (0,\,1)$, we get for the whole expression that $\partial V_i(\mathbf{p},\omega)/\partial p_j|_{\partial V_i(\mathbf{p},\omega)/\partial p_i = 0} \in (0,\,1)$, satisfying Condition 1. □

Proof of Proposition 3

Proof. First, note that in $t = 1$ we have $W_{i,j} = 0$ for all $j \in \{0, 1, ..., n\}$. Second, note that if $\min_i x_i \geq t$, $i \in J(\boldsymbol{\omega})$, all firms' continuation values are equal for any consumer choice, i.e. $W_{i,j}(\boldsymbol{\omega}) = W_{i,k}(\boldsymbol{\omega})$ for all $j, k \in J(\boldsymbol{\omega}) \cup \{0\}$ and all $i \in J(\boldsymbol{\omega})$. To see this consider the following backward induction argument. In $t = 2$, each active firm holds sufficient capacity to sell one good now and one good in $t = 1$. Consequently, independent of the consumer choice j in $t = 2$, in the final period $t = 1$ all firms will be competing, such that the value function of all firms i in $t = 1$, as expected in $t = 2$, i.e. $W_{i,j}(\boldsymbol{\omega})$, will be identical for all $j \in J(\boldsymbol{\omega}) \cup \{0\}$. Similarly, this holds in any t where each active firm holds sufficient capacity for all periods until the deadline. Therefore, all demand realizations yield equal continuation values (of harsh competition).

Hence, here the first-order condition from equation (12) simplifies to

$$\mu = \sum_{j \neq i} D_j\Big(\mathbf{p}, \boldsymbol{\omega}\Big)\, p_i \Longleftrightarrow \frac{p_i}{\mu} = \frac{1}{\sum\limits_{j \neq i} D_j\Big(\mathbf{p}, \boldsymbol{\omega}\Big)} \geq 1. \tag{A.9}$$

The inequality holds because $\sum_{j \neq i} D_j\left(\mathbf{p}, \boldsymbol{\omega}\right) \leq 1$. Similarly, here (A.7) simplifies to

$$\frac{\partial V_i(\mathbf{p}, \boldsymbol{\omega})}{\partial p_j}\Big|_{\frac{\partial V_i(\mathbf{p},\omega)}{\partial p_i} = 0} = D_j(\mathbf{p}, \boldsymbol{\omega}) \left[\frac{p_i}{\mu} - 1\right] \geq 0. \tag{A.10}$$

Note that this inequality is because $p_i/\mu \geq 1$ by (A.9). Further, if we insert p_i/μ from (A.9) into (A.10) we get

$$\frac{\partial V_i(\mathbf{p}, \boldsymbol{\omega})}{\partial p_j}\Bigg|_{\frac{\partial V_i(\mathbf{p},\omega)}{\partial p_i} = 0} = \frac{D_j(\mathbf{p}, \boldsymbol{\omega})}{D_j(\mathbf{p}, \boldsymbol{\omega}) + \sum\limits_{k \neq i,j} D_k\Big(\mathbf{p}, \boldsymbol{\omega}\Big)} - D_j(\mathbf{p}, \boldsymbol{\omega}) < 1. \tag{A.11}$$

This inequality holds because the first term cannot be larger than one. From equations (A.10) and (A.11) we see that Condition 1 holds and the equilibrium is unique if $t = 1$ or $t > 1$ with $\min_i x_i \geq t$ and $i \in J(\boldsymbol{\omega})$. □

A.2 Forward-looking Consumers

Definition of consumer transition matrix

Let $\Phi(\boldsymbol{\omega})$ be initially a zero-matrix of dimension $(T \times \boldsymbol{X} \times C)$. Then, each entry $(t', \mathbf{x}', c') \in \{\{0, .., T\} \times \{0, .., \boldsymbol{X}\} \times \{1, .., C\}\}$, where $t' < t$, $x'_j \leq x_j$, $\forall j$, is recursively defined by $\Phi(\boldsymbol{\omega})_{[t', \mathbf{x}', c']}$

$$
= \begin{cases}
\Bigg[\sum\limits_{j \in J(\boldsymbol{\omega})} \Big\{ D_j \left(t'+1, x'_j+1, \boldsymbol{x'_{-j}}, c' \right) \Big\} \frac{c'-1}{c'} \\
\quad + D_0 \left(t'+1, \boldsymbol{x'}, c'-1 \right) \frac{c'-2}{c'-1} \Bigg] \frac{1}{c'} & \text{if } t' = t-1 > 0, \\
\Bigg[\sum\limits_{j \in J(\boldsymbol{\omega})} \Big\{ D_j \left(t'+1, x'_j+1, \boldsymbol{x'_{-j}}, c' \right) \Big\} \frac{c'}{c'+1} \\
\quad + D_0 \left(t'+1, \boldsymbol{x'}, c'-1 \right) \frac{c'-2}{c'-1} \Bigg] & \text{if } t' = t-1 = 0, \\
\Bigg[\sum\limits_{j \in J(\boldsymbol{\omega})} \Big\{ \Phi(\boldsymbol{\omega})_{\left[t'+1, x'_j+1, \boldsymbol{x'_{-j}}, c'\right]} \, c' \, D_j \left(t'+1, x'_j+1, \boldsymbol{x'_{-j}}, c' \right) \Big\} \frac{c'-1}{c'} \\
\quad + \Phi(\boldsymbol{\omega})_{[t'+1, \boldsymbol{x'}, c'-1]} \left(c'-1 \right) D_0 \left(t'+1, \boldsymbol{x'}, c'-1 \right) \frac{c'-2}{c'-1} \Bigg] \frac{1}{c'} & \text{if } t-1 > t' > 0, \\
\Bigg[\sum\limits_{j \in J(\boldsymbol{\omega})} \Big\{ \Phi(\boldsymbol{\omega})_{\left[t'+1, x'_j+1, \boldsymbol{x'_{-j}}, c'\right]} \, c' \, D_j \left(t'+1, x'_j+1, \boldsymbol{x'_{-j}}, c' \right) \Big\} \frac{c'-1}{c'} \\
\quad + \Phi(\boldsymbol{\omega})_{[t'+1, \boldsymbol{x'}, c'-1]} \left(c'-1 \right) D_0 \left(t'+1, \boldsymbol{x'}, c'-1 \right) \frac{c'-2}{c'-1} \Bigg] & \text{if } t-1 > t' = 0.
\end{cases} \tag{A.12}
$$

For the completion of the recursion note that for all states with $t' \geq t$ or with $x'_j > x_j$ for some $j \in J(\boldsymbol{\omega})$, we have that $\Phi(\boldsymbol{\omega})_{[t', \mathbf{x}', c']} = 0$ because these states cannot be reached as they either lie in the past, are equal to the current time period or consist of infeasible capacities. Further note that the matrix contains all probabilities to reach the deadline without being lucky again, too, hence the sum over all its entries must add up to one.

As an example of how to read $\Phi(\boldsymbol{\omega})_{[t', \mathbf{x}', c']}$ consider the third case. The state $[t', \mathbf{x}', c']$ will be reached after two possibilities. First, after the sum of probabilities D_j that in state $(t'+1, x'_j+1, \boldsymbol{x'_{-j}}, c')$, which itself will have been reached with probability $\Phi(\boldsymbol{\omega})_{\left[t'+1, x'_j+1, \boldsymbol{x'_{-j}}, c'\right]} \, c'$, some firm $j \in J(\boldsymbol{\omega})$ will have sold a good to some other consumer. This means that the currently lucky consumer will not have been lucky then, s.t. this will have been weighted with probability $(c'-1)/c'$.[69] Second, after the lucky consumer of state $(t'+1, \boldsymbol{x'}, c'-1)$, which will have been reached with probability $\Phi(\boldsymbol{\omega})_{[t'+1, \boldsymbol{x'}, c-1']} \, (c'-1)$, will have chosen to wait, i.e. probability D_0. Again, this

[69]No state with $c' = 1$ can be reached, i.e. is assigned a probability of zero, because after new consumer arrival there will be at least one more consumer, next to the currently lucky one.

means that the currently lucky consumer will not have been lucky then, s.t. this will have been weighted with probability $(c'-2)/(c'-1)$.[70] Finally, the whole expression is weighted with $1/c'$, i.e. the probability to be lucky in $(t', \mathbf{x}', c')$.[71]

Proof of Proposition 4

The proof is identical to the ones for Lemmas 1 and 2 as well as Proposition 1. The only technical difference is that for the upper bound of the best-response function since

$$1 - D_i(\mathbf{p}, \boldsymbol{\omega}) = 1 - \frac{\exp(\frac{v-p_i}{\mu})}{\exp[\delta_c W_c(\boldsymbol{\omega})] + \sum\limits_{J(\boldsymbol{\omega})} \exp(\frac{v-p_j}{\mu})} \geq \frac{1}{1 + \left(\exp(\frac{v}{\mu}) / \exp[\delta_c W_c(\boldsymbol{\omega})]\right)}$$

we have that $\sup p_i(\mathbf{p_{-i}}, \boldsymbol{\omega}) \leq \mu\{1 + (\exp(\frac{v}{\mu}) / \exp[\delta_c W_c(\boldsymbol{\omega})])\} + \delta w_i^+(\boldsymbol{\omega}) < \infty$.

Proof of Proposition 5

This proof is identical to the ones for Propositions 2 and 3. Note only that for $\mu \to 0$ also utility of waiting $U_0 \to 0$. Then demand for waiting $D_0(\mathbf{p}, \boldsymbol{\omega}) \to 0$, too.

Derivatives of demand w.r.t. δ_c in $t = 2$

In $t = 2$ all expected continuation values $W_c(\boldsymbol{\omega})$ are independent of δ_c because after $t = 1$ there is only the deadline. First, abbreviate the denominator of demand by $N := \exp[\delta_c W_c] + \sum_{J(\boldsymbol{\omega})} \exp\left(\frac{v-p_j}{\mu}\right)$. Then, for all $i, j \in J(\boldsymbol{\omega})$ and $j \neq i$ we have

$$\frac{\partial D_i(\mathbf{p})}{\partial \delta_c} = \frac{0 - \left[\exp\left(\frac{v-p_i}{\mu}\right)\right][\exp[\delta_c W_c] W_c]}{[N]^2} = -W_c(\boldsymbol{\omega}) D_i(\mathbf{p}) D_0(\mathbf{p}) < 0$$

$$\frac{\partial D_0(\mathbf{p})}{\partial \delta_c} = \frac{N \exp[\delta_c W_c] W_c - \exp[\delta_c W_c] \exp[\delta_c W_c] W_c}{[N]^2} = W_c(\boldsymbol{\omega}) D_0(\mathbf{p}) (1 - D_0(\mathbf{p})) > 0$$

$$\partial[\sum_{j \neq i} D_j(\boldsymbol{p}) + D_0(\boldsymbol{p})] / \partial \delta_c = \partial [1 - D_i(\boldsymbol{p})] / \partial \delta_c \quad > \quad 0$$

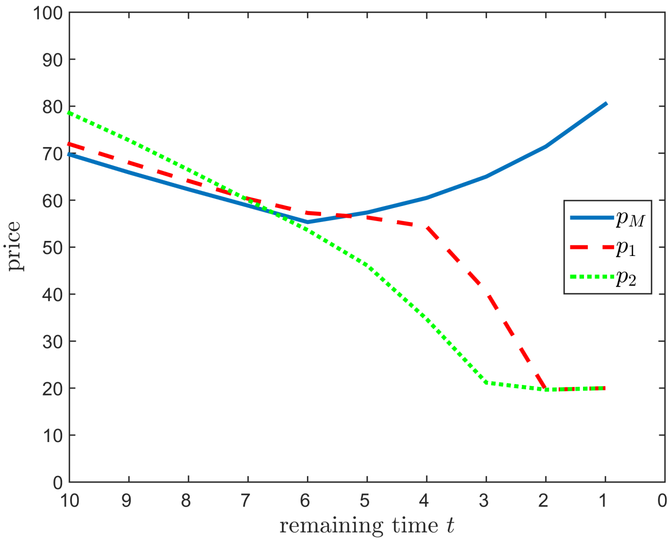
[70] $\Phi(\boldsymbol{\omega})$ gives the probabilities of when the lucky consumer will be for the first time again the lucky consumer, i.e. no further continuations need to be considered.

[71] Note for the cases with $t' = 0$ that the expression is weighted with 1 as this would be the deadline reached by all remaining consumers.

A.3 Comparative Statics and Policy

Comparative Statics: Capacity Cases

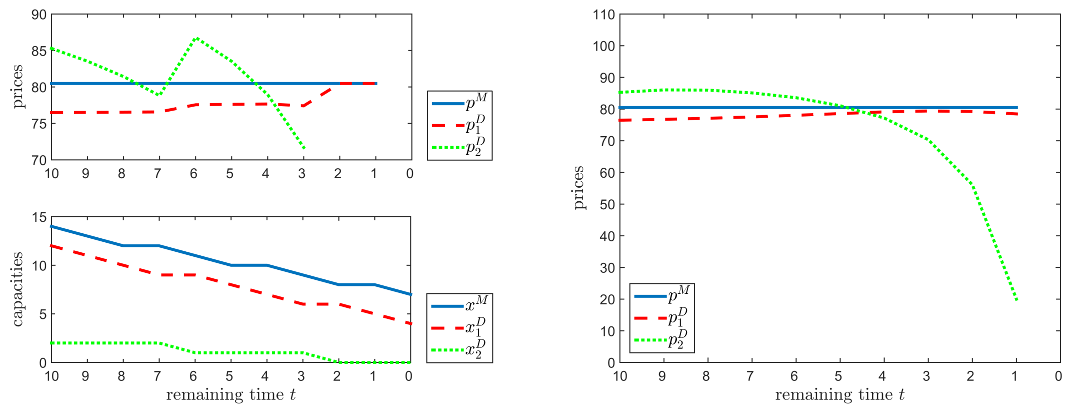
FIGURE A.1: Prices with Fixed Capacities (F-l Consumers)



Parameters: $T = 10$, $v = 100$, $\delta = 1$, $\mu = 10$, forward-looking consumers. Capacities are fixed, s.t. under duopoly $x_1 = 4$, $x_2 = 2$ and under monopoly $x_M = 6$ in all t.

FIGURE A.2: Asymmetric Aggregate Excess Capacities (2)

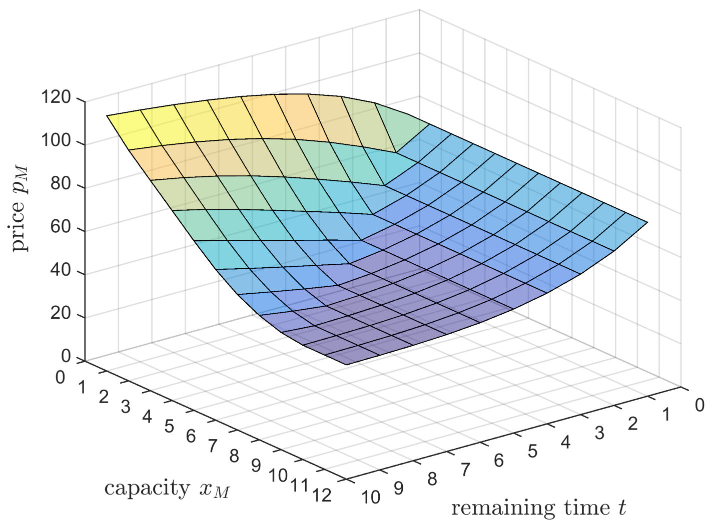
(a) Representative Price and Capacity Path (b) Average Price Path



Parameters: $T = 10$, $v = 100$, $\delta = 1$, $\mu = 10$, myopic consumers. Initial capacities under duopoly are $x_1 = 12$, $x_2 = 2$ and under monopoly $x_M = 14$.

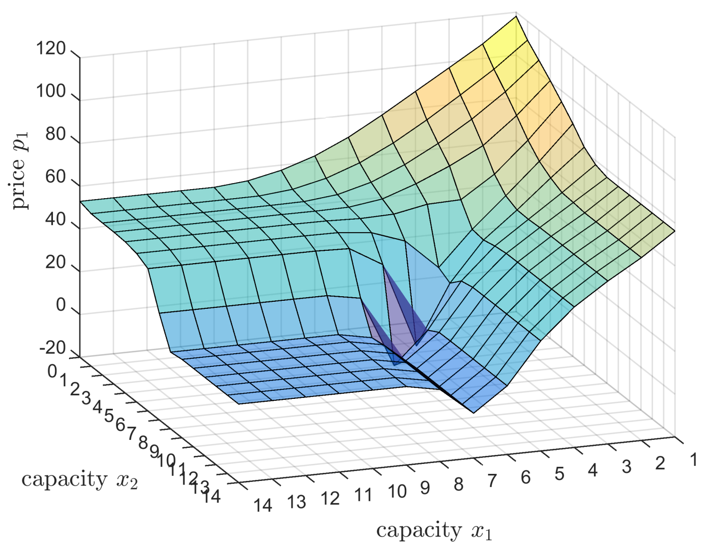
Comparative Statics: Capacity Levels and Remaining Time (F-l Consumers)

FIGURE A.3: Comp. Statics of Capacity and Time (Monopoly, F-l Cons.)



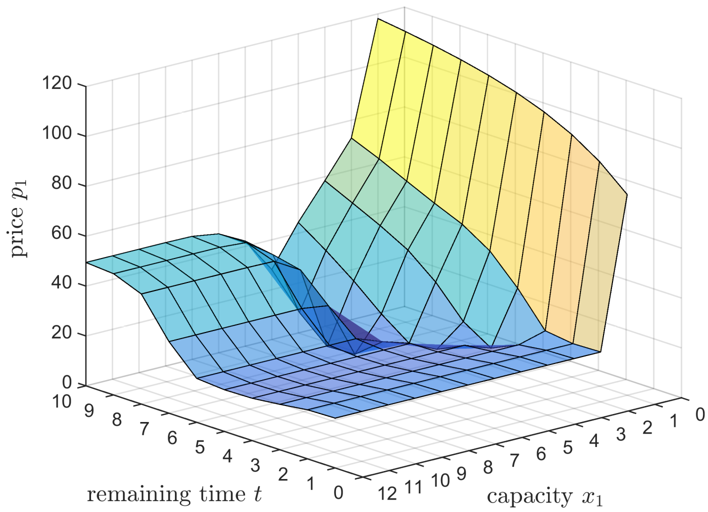
Parameters: $v = 100$, $\delta = 1$, $\mu = 10$, forward-looking consumers. Monopoly price p_M for different capacity levels of x_M and different remaining time periods t.

FIGURE A.4: Comp. Statics of Capacity Levels (Duopoly, F-l Cons.)



Parameters: $t = 10$, $v = 100$, $\delta = 1$, $\mu = 10$, forward-looking consumers. Price p_1 in $t = 10$ under duopoly for different capacity levels of x_1 and x_2.

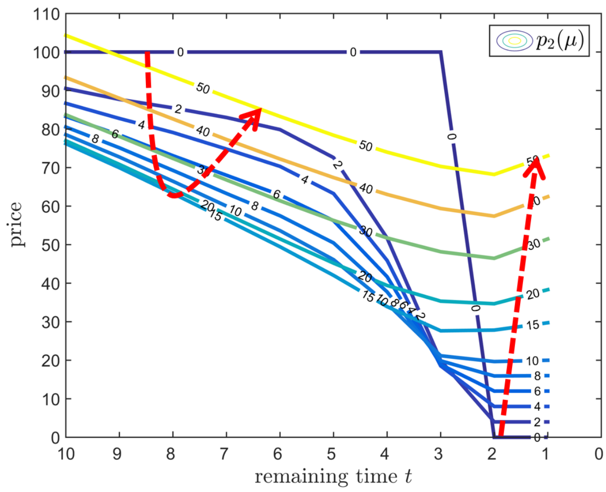
FIGURE A.5: Comp. Statics of Remaining Time (Duopoly, F-l Cons.)



Parameters: $v = 100$, $\delta = 1$, $\mu = 10$, forward-looking consumers. Price p_1 under duopoly for different levels of capacity x_1 and different remaining time periods t, while $x_2 = 5$ is fixed.

Comparative Statics: Consumer Heterogeneity

FIGURE A.6: Comp. Statics of Consumer Heterogeneity (F-l Cons.)



Parameters: $T = 10$, $v = 100$, $\delta = 1$, forward-looking consumers. Capacities are fixed, s.t. $x_1 = 4$, $x_2 = 2$ in all t. The price $p_2(\mu)$ is for different levels of consumer heterogeneity $\mu \in \{0, 2, 4, 6, 8, 10, 15, 20, 30, 40, 50\}$, where the "in-line" number refers to μ. The dashed arrows show the effect of μ on p_2, one for the region $x_2 = 2 < t$ and the other for the region $x_2 = 2 \geq t$.

Proof of Lemma 3

Proof. In period $t = 1$ there is no continuation value for consumers (nor firms) and hence demand choices and equilibrium prices are identical for all levels of δ_c. Therefore in $t = 1$ also the valuation function $V_c(\boldsymbol{\omega})$ of consumers will be independent of δ_c. Hence, all firm continuation valuations $W_{i,j}(\boldsymbol{\omega})$, for all i, j, are independent of δ_c if $t \leq 2$. Consequently in $t = 2$, the only part of the demand function in (18) depending on δ_c will be the actual discount factor in the term $\exp[\delta_c W_c(\boldsymbol{\omega})]$. For the effect on equilibrium prices in $t = 2$ recall the first-order condition in equation (12), which is analogous for the forward-looking consumer case. Since the sum of all but i's demands increase in δ_c (compare Appendix A.2), equilibrium price p_i^* must decrease in δ_c. Only if capacities are individually excessive, all continuations are equal for the firms, such that the effect of δ_c on demands cancels out. Overall we have $\partial p_i^* / \partial \delta_c \leq 0$ for $t \leq 2$. □

Proof of Lemma 4

Proof. In the last period, $t = 1$, if there is no consumer, the continuation value for firms is zero. If, however, a consumer arrives, pricing is not affected by λ because there are no future arrivals. Then, in the pen-ultimate period, $t = 2$, if a consumer arrives (otherwise the problem is irrelevant) the maximization problem of the firms is as given at equation (6), with the only difference that the transition matrix is now weighted with the probability of consumer arrival in $t = 1$, i.e.

$$\boldsymbol{\omega'} | (\boldsymbol{\omega}, j) = \begin{cases} (t-1, \mathbf{x}, c=1) & \text{if } j = 0, \\ (t-1, \mathbf{x} - \mathbf{e}_j, c=1) & \text{if } j \in J(\boldsymbol{\omega}), \end{cases} \quad \text{with probability } \lambda,$$

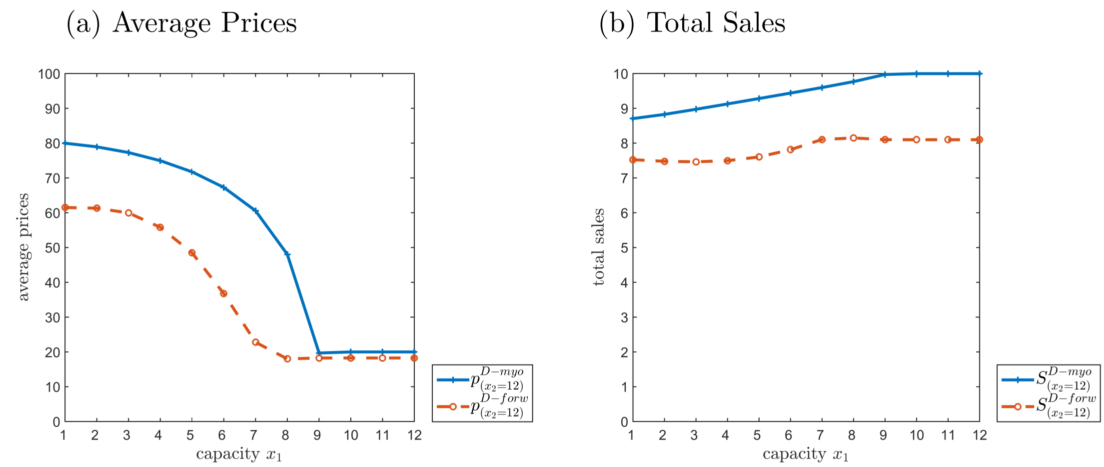
$$\boldsymbol{\omega'} | (\boldsymbol{\omega}, j) = \begin{cases} (t-1, \mathbf{x}, c=0) & \text{if } j = 0, \\ (t-1, \mathbf{x} - \mathbf{e}_j, c=0) & \text{if } j \in J(\boldsymbol{\omega}), \end{cases} \quad \text{with probability } 1 - \lambda.$$

Note that the state $\boldsymbol{\omega}$ now also describes the number of consumers c, as in the model with forward looking-consumers, although with myopic consumers we can only have one or no consumer in the market. The expected continuation value increases in λ because in the case of no arrival $(1 - \lambda)$ there will be no demand and hence zero revenues, whereas in the case of consumer arrival (λ) the continuation value is identical to the standard model with $\lambda = 1$. Now consider the resulting equilibrium condition in $t = 2$, as given by equation (12) and note for the continuation values that $\delta W_{i,i}(\boldsymbol{\omega}) - \delta W_{i,j}(\boldsymbol{\omega}) \leq 0$ (for all $j \in J(\boldsymbol{\omega}) \cup \{0\}$, $j \neq i$) because either firm i will have sold out and hence $W_{i,i}(\boldsymbol{\omega}) = 0$

or competitors j will have sold out, hence i remains monopolist and $W_{i,i}(\boldsymbol{\omega}) < W_{i,j}(\boldsymbol{\omega})$, where $i, j \in J(\boldsymbol{\omega})$, or all firms remain with positive capacity after any sale and then $W_{i,i}(\boldsymbol{\omega}) = W_{i,j}(\boldsymbol{\omega})$, while always $W_{i,i}(\boldsymbol{\omega}) \leq W_{i,0}(\boldsymbol{\omega})$ because any positive capacity in $t = 1$ allows for at most one unit to be sold. As argued above, all continuation values are "discounted" symmetrically with weight $(1-\lambda)$.[72] Consequently $\delta W_{i,i}(\boldsymbol{\omega}) - \delta W_{i,j}(\boldsymbol{\omega})$ weakly increases (becomes less negative) for all $j \in J(\boldsymbol{\omega}) \cup \{0\}$, $j \neq i$, in λ and the equilibrium price must weakly increase in λ. $\square$

Welfare Analysis: Forward-looking vs Myopic Consumers

FIGURE A.7: Forward-looking vs Myopic Consumers: Welfare (2)



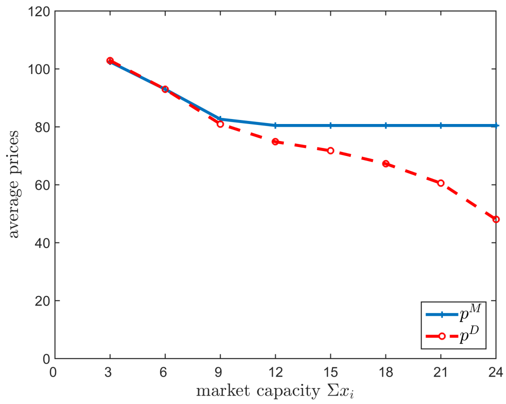
Parameters: $t = 10$, $v = 100$, $\delta = 1$, $\mu = 10$. Averages prices (a) total sales (b), as expected in $t = 10$, under monopoly and duopoly for different levels of initial capacity x_1, given fixed levels for $x_2 = 12$, for myopic (myo: $\delta_c = 0$) as well as forward-looking consumers (forw: $\delta_c = \delta$).

[72] If consumers are forward-looking, the problem is only relevant if only one consumer is in the market, hence the proof is analogous.

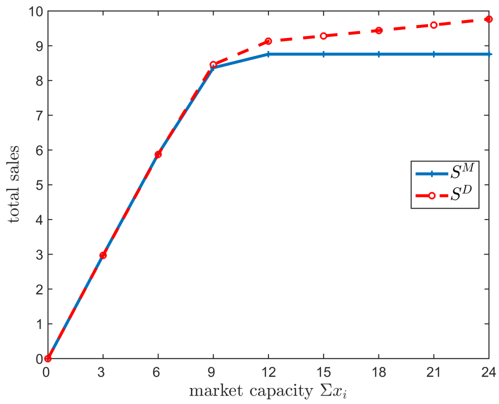
Welfare Analysis: Competition Policy

FIGURE A.8: Competition Policy (2)

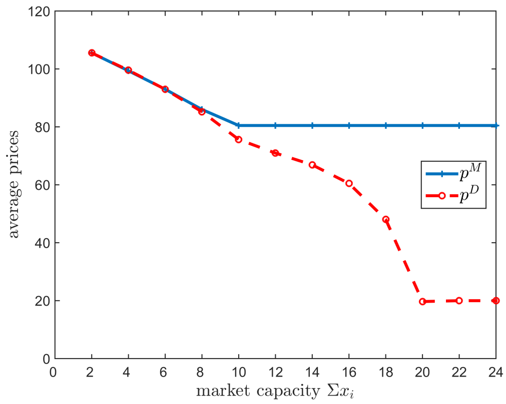
(a) Average Prices (extreme asymmetry)



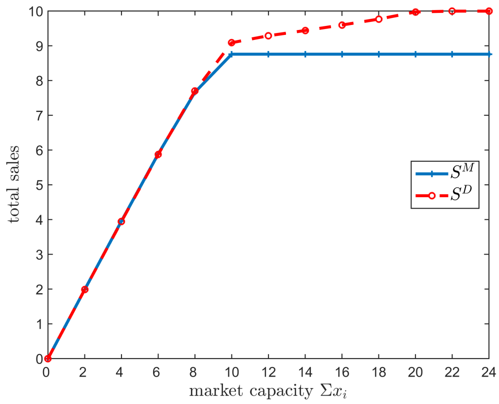
(b) Total Sales (extreme asymmetry)



(c) Average Prices (simple asymmetry)



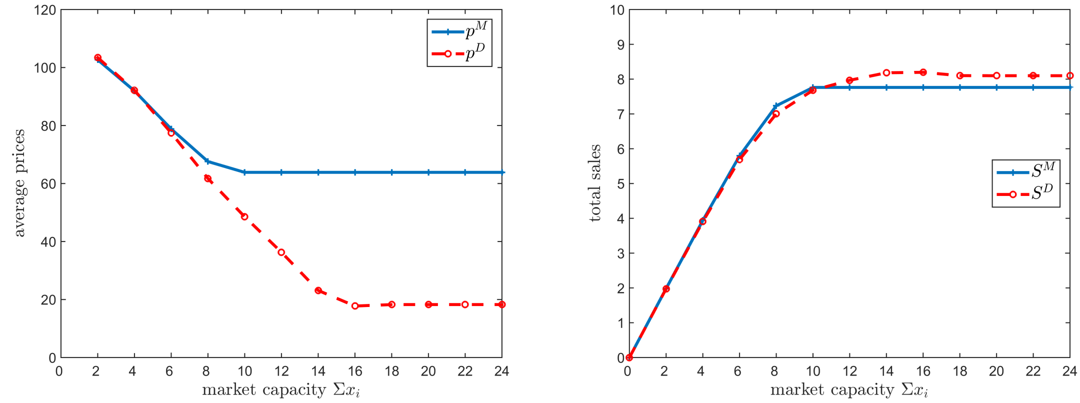
(d) Total Sales (simple asymmetry)



Parameters: $t = 10$, $v = 100$, $\delta = 1$, $\mu = 10$, myopic consumers. Average prices and total sales as expected in $t = 10$, under monopoly and duopoly for different levels of initial market capacity $\sum_i x_i$
(a) and (b): extremely asymmetric duopoly capacities, such that $x_1^D = 2x_2^D = x^M/3$,
(c) and (d): less extremely asymmetric duopoly capacities, such that $x_1^D = x_2^D + 2 = x^M/2 + 1$,
while only if $\sum_i x_i = 2$, let $x_1^D = x_2^D$.

FIGURE A.9: Competition Policy (F-l Consumers)

(a) Average Prices (b) Total Sales

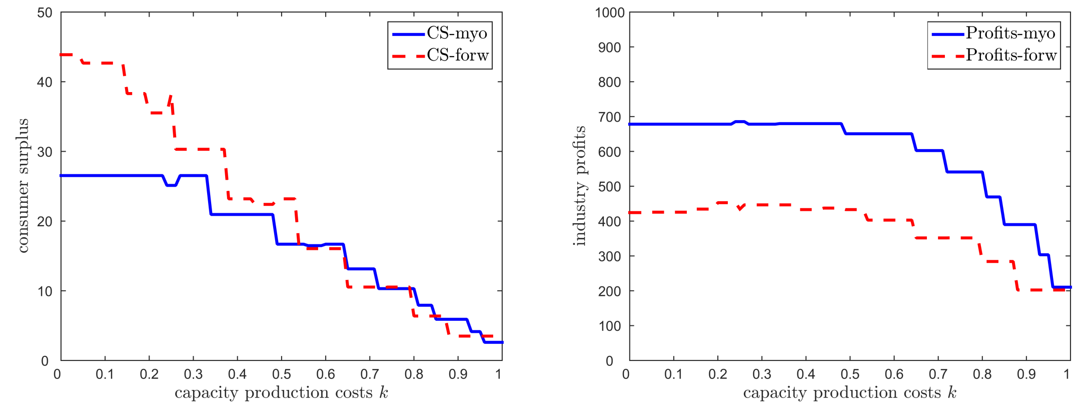


Parameters: $t = 10$, $v = 100$, $\delta = 1$, $\mu = 10$, forward-looking consumers. Average prices (a) and total sales (b) as expected in $t = 10$, under monopoly and (initially symmetric) duopoly for different levels of initial market capacity $\sum_i x_i$, s.t. $x_1^D = x_2^D = x^M/2$ in $t = 10$.

Welfare Analysis: Ex-ante Capacity Production

FIGURE A.10: Welfare Given the Ex-ante Capacity Equilibrium (2)

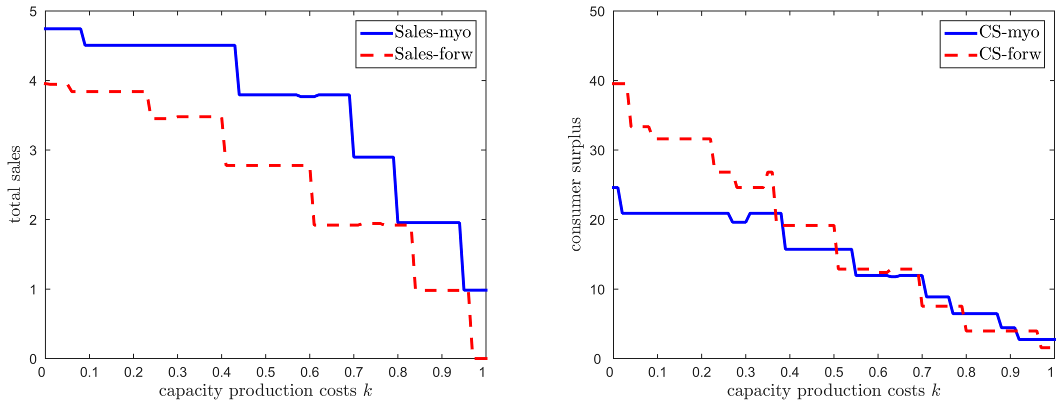
(a) Consumer Surplus (b) Industry Profits



Parameters: $T = 10$, $v = 100$, $\delta = 1$, $\mu = 10$, duopoly for myopic consumers and forward-looking consumers with $\delta_c = \delta = 1$. Consumer surplus (a) and industry profits (b) of the dynamic pricing game after ex-ante equilibrium capacity production, for different k. For multiple equilibria I consider the mean of their pure-strategies.

FIGURE A.11: Welfare Given the Ex-ante Capacity Equilibrium (3)

(a) Total Sales, $T = 5$ (b) Consumer Surplus, $T = 8$



Parameters: $v = 100$, $\delta = 1$, $\mu = 10$, duopoly for myopic consumers and forward-looking consumers with $\delta_c = \delta = 1$. Total sales in $T = 5$ (a) and consumer surplus in $T = 8$ (b) of the dynamic pricing game after ex-ante equilibrium capacity production, for different k. For multiple equilibria I consider the mean of their pure-strategies.

B Algorithm for Equilibrium (Programming Code)

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%
% Dynamic Pricing under Capacity-constrained Competition with a Deadline
% Philipp Dimakopoulos, November 2017
%
% Duopoly Simulation, Forward-Looking Consumers
% Based on elements from Ulrich Doraszelski's BDPEMS Short Course, 2015
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% 1) PREPARATION

% clear workspace and create diary
clear;
delete('duopoly.lst');
diary('duopoly.lst');
```

Global Variables

```
global beta A B C T H Q mc v mm arr mufix;
% A: max number of firm A capacity, a=1 means no capacity
% B: max number of firm B capacity, b=1 means no capacity
% T-1: total selling periods, t = T is the deadline
% C: max number of consumers in the market, c=1 means no consumer
% H: the number of possible mu - values
% Q: the number of possible betac - values
% mc: marginal costs (=0)
% v: baseline value of a product
% mm: grid for global deviation detection
% arr: arrival rate of consumers
% mufix: fix value for mu

% Run setup program giving inital capacities and remaining time, etc.
run setup_s;

% define empty value matrices, for iterations
V00A = zeros(T,A,B,C,H,Q);
V00B = zeros(T,A,B,C,H,Q);
V0A = V00A;
V0B = V00B;
V1A = V00A;
V1B = V00B;
V00C = zeros(T,A,B,C,H,Q);
V0C = V00C;
V1C = V00C;

% define demand matrices
De0 = zeros(T,A,B,C,H,Q);
DeA = zeros(T,A,B,C,H,Q);
DeB = zeros(T,A,B,C,H,Q);

% define price matrices
pA = zeros(T,A,B,C,H,Q);
```

```
pB = zeros(T,A,B,C,H,Q);
pA_0 = pA;
pB_0 = pB;
pA_0_G = pA;
pB_0_G = pB;

% define welfare measures
SumDn0W = zeros(T,A,B,C,H,Q);   % sum of total sales, "efficiency"
CS = zeros(T,A,B,C,H,Q);        % consumer surplus
ProfA = zeros(T,A,B,C,H,Q);
ProfB = zeros(T,A,B,C,H,Q);
Profs = zeros(T,A,B,C,H,Q);

% starting values for done2 loop
startval1 = lowerstart;
startval2 = lowerstart;
pA00 = startval1.*(v/10).*ones(T,A,B,C,H);
pB00 = startval2.*(v/10).*ones(T,A,B,C,H);

% time of calculation
tic

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% 2) ITERATION

disp(' ');
disp('Value function iteration');
disp('------------------------');
disp(' ');
```

Loop for each state, beginning in the end

```
for t=T:-1:1
    % t=T is the deadline, where no sales take place
    if t==T
        for c=1:C
        for h=1:H
        for a=1:A
        for b=1:B
        for q=1:Q
             % continuation values of firms and consumers = 0
            V0A(t,a,b,c,h,q)=0;
            V0B(t,a,b,c,h,q)=0;
            V0C(t,a,b,c,h,q)=0;
            % infinite prices s.t. no demand
            pA(t,a,b,c,h,q)=realmax;
            pB(t,a,b,c,h,q)=realmax;
        end
        end
        end
        end
        end
    % for all t<T
    else
```

```
        for c=1:(t+1)    % there can never be more than t+1 consumers
        for h=1:H
        for a=1:A
        for b=1:B
        for q=1:Q
```

done2-Loop over different start values for root-finding

```
            iter2 = 0;
            done2 = 0;
            while ~done2

% for done2 loop: new start values
pA_0(t,a,b,c,h) = startval1.*(v/10);
pB_0(t,a,b,c,h) = startval2.*(v/10);
pA_0_G(t,a,b,c,h) = startval1.*(v/10);
pB_0_G(t,a,b,c,h) = startval2.*(v/10);

% consumer heterogeneity from function mufun
mu = mufun(h);
% consumer patience from function betacfun
betac = betacfun(q);

% transition probabilties for firms:
% input: state of the world
% output: state reached after j in (wait, buyA, buyB)
% use transitionprob function
[PrL_0,PrW_A,PrW_B] = transprobs_st_s(t,a,b,c,h,q);

% transition probabilties for lucky consumer
% input: state of the world
% output: state reached next time as lucky consumer if j=0
% use transitionprob function
[PrC_0] = transprobs_st_s_cons(t,a,b,c,h,q,De0,DeA,DeB);

% check
disp(sprintf('%2d %2d %2d %2d %2d %2d',t,a,b,c,h,q));

% Value for i in {A and B}, conditional on j [notation here: Wj_i]
W0_A = sum(sum(sum(sum(sum(sum(PrL_0.*V1A))))));
W0_B = sum(sum(sum(sum(sum(sum(PrL_0.*V1B))))));

WA_A = sum(sum(sum(sum(sum(sum(PrW_A.*V1A))))));
WA_B = sum(sum(sum(sum(sum(sum(PrW_A.*V1B))))));

WB_A = sum(sum(sum(sum(sum(sum(PrW_B.*V1A))))));
WB_B = sum(sum(sum(sum(sum(sum(PrW_B.*V1B))))));

%  Value for lucky consumer, conditional on j=0
WW_C = sum(sum(sum(sum(sum(sum(PrC_0.*V1C))))));
```

doneGlob-Loop: global deviations?

```
doneGlob = 0;
iterGlob = 0;
```

```
while ~doneGlob  %global price max

    % help price matrices
    pA_0(t,a,b,c,h,q) = pA_0_G(t,a,b,c,h,q);
    pB_0(t,a,b,c,h,q) = pB_0_G(t,a,b,c,h,q);
```

done-Loop: internal price quilibrium

```
    done = 0;
    iter = 0;
    while ~done

% parameters for maximization
MaxIter = 400; TolFun = 1e-6; TolX = 1e-6;
options = [1,TolFun,TolX,TolFun,MaxIter,1e-6];

% Equilibrium for different cases
% using FOC functions,
% which uses demand functions
% and Root-Finding-Algorithm (DogLeg)

% no capas or no consumer
if ( a == 1 && b == 1 ) || c==1
    pA(t,a,b,c,h,q) = realmax;
    pB(t,a,b,c,h,q) = realmax;
    info(6)=0;
    infoB(6)=0;

% only B has capa
elseif a == 1 && b~=1
    pA(t,a,b,c,h,q) = realmax;
    [pB(t,a,b,c,h,q),infoB] = ...
        DogLeg('FOC_st_s_B', ...
        [mc,pA_0(t,a,b,c,h,q),W0_B,WA_B,WB_B,WW_C,a,b,h,q], ...
        pB_0(t,a,b,c,h,q),options);
    info(6)=0;

% only A has capa
elseif b == 1 && a~=1
    pB(t,a,b,c,h,q) = realmax;
    [pA(t,a,b,c,h,q),info] = ...
        DogLeg('FOC_st_s_A', ...
        [mc,pB_0(t,a,b,c,h,q),W0_A,WA_A,WB_A,WW_C,a,b,h,q], ...
        pA_0(t,a,b,c,h,q),options);
    infoB(6)=0;

% both have capa: dual maximization
else
    pDog_0 = zeros(2,1);
    pDog_0(1) = pA_0(t,a,b,c,h,q);
    pDog_0(2) = pB_0(t,a,b,c,h,q);
    [pDog,info] = ...
        DogLeg('FOC_st_s_AB', ...
        [mc,mc,W0_A,W0_B,WA_A,WA_B,WB_A,WB_B,WW_C,a,b,h,q], ...
```

```
        pDog_0,options);
    pA(t,a,b,c,h,q) = pDog(1);
    pB(t,a,b,c,h,q) = pDog(2);
end

% report possible error
if info(6)>3 || infoB(6)>3
    warning('There is a problem with DogLeg!');
    pause
end

% internal EQ loop ok (done = 1) if tolerance ok
tolpA = abs(pA(t,a,b,c,h,q)-pA_0(t,a,b,c,h,q))/(1+abs(pA(t,a,b,c,h,q)));
tolpB = abs(pB(t,a,b,c,h,q)-pB_0(t,a,b,c,h,q))/(1+abs(pB(t,a,b,c,h,q)));
if ((tolpA<tol) && (tolpB<tol)) || (iter>=maxiter)
    done = 1;
end

% if tolerance level not reached -> another EQ loop
if ~done
    iter = iter + 1;
end
% display iteration numbers
if (mod(iter,10)==0) || done
    disp(sprintf('iter=%d tolpA=%g tolpB=%g',iter,tolpA,tolpB));
end
% old price as new starting value
pA_0(t,a,b,c,h,q) = pA(t,a,b,c,h,q);
pB_0(t,a,b,c,h,q) = pB(t,a,b,c,h,q);

    end     % end done of internal price equilibrium

    % Update value function.
    % get demand for EQ prices
    [D0,DA,DB] = demand_st_s(pA(t,a,b,c,h,q),pB(t,a,b,c,h,q),WW_C,a,b,h,q);

    De0(t,a,b,c,h,q) = D0;
    DeA(t,a,b,c,h,q) = DA;
    DeB(t,a,b,c,h,q) = DB;

    % firms
    V1A(t,a,b,c,h,q) = DA.*(pA(t,a,b,c,h,q)-mc) ...
                       + beta.*(D0.*W0_A+DA.*WA_A+DB.*WB_A) ;
    V0A(t,a,b,c,h,q) = V1A(t,a,b,c,h,q);

    V1B(t,a,b,c,h,q) = DB.*(pB(t,a,b,c,h,q)-mc) ...
                       + beta.*(D0.*W0_B+DA.*WA_B+DB.*WB_B);
    V0B(t,a,b,c,h,q) = V1B(t,a,b,c,h,q);

    % consumers
    uw = exp( betac.*(WW_C) );
    if ( a==1 && b == 1 ) || c==1
        V1C(t,a,b,c,h,q) = 0 ;
    elseif a==1 && b ~= 1
```

```
        V1C(t,a,b,c,h,q) ...
        = log( 0 + exp( (v-pB(t,a,b,c,h,q))./mu ) + uw );
    elseif b==1 && a ~= 1
        V1C(t,a,b,c,h,q) = ...
        log( exp( (v-pA(t,a,b,c,h,q))./mu ) + 0 + uw );
    else
        V1C(t,a,b,c,h,q) = ...
        log( (1/2)* ( exp( (v-pA(t,a,b,c,h,q))./mu ) ...
        + exp( (v-pB(t,a,b,c,h,q))./mu ) + 2*uw ) );
    end
    V0C(t,a,b,c,h,q) = V1C(t,a,b,c,h,q);
```

Define welfare measures. To be adjusted for arrival rates, if neccessary!

```
    if t == T
        SumDn0W(t,a,b,c,h,q) = 0;
        WF(t,a,b,c,h,q) = 0;
        ProfA(t,a,b,c,h,q) = 0;
        ProfB(t,a,b,c,h,q) = 0;
        Profs(t,a,b,c,h,q) = 0;
        CS(t,a,b,c,h,q) = 0;
        Pricing(t,a,b,c,h) = NaN;
    else
        [PrL_0,PrW_A,PrW_B] = transprobs_st_s(t,a,b,c,h,q);
        Sum_0 = sum(sum(sum(sum(sum(sum(PrL_0.*SumDn0W))))));
        Sum_A = sum(sum(sum(sum(sum(sum(PrW_A.*SumDn0W))))));
        Sum_B = sum(sum(sum(sum(sum(sum(PrW_B.*SumDn0W))))));
        CS_0 = sum(sum(sum(sum(sum(sum(PrL_0.*CS))))));
        CS_A = sum(sum(sum(sum(sum(sum(PrW_A.*CS))))));
        CS_B = sum(sum(sum(sum(sum(sum(PrW_B.*CS))))));

        % sum of realized trades = efficiency
        SumDn0W(t,a,b,c,h,q) = DA + DB ...
           + beta.*( D0.*Sum_0 + DA.*Sum_A + DB.*Sum_B );

        % consumer surplus
        uw = exp( betac.*(WW_C) );
        if a==1 && b == 1 && c>0
            CS(t,a,b,c,h,q) = 0;
        elseif c==1
            CS(t,a,b,c,h,q) = ...
                beta.*( D0.*CS_0 + DA.*CS_A + DB.*CS_B);
        elseif a==1 && b ~= 1 && c>1
            CS(t,a,b,c,h,q) = ...
                log( 1 + 0 + exp( (v-pB(t,a,b,c,h,q))./mu ) ) ...
                + beta.*( D0.*CS_0 + DA.*CS_A + DB.*CS_B);
        elseif a~=1 && b == 1 && c>1
            CS(t,a,b,c,h,q) = ...
                log( 1 + exp( (v-pA(t,a,b,c,h,q))./mu ) + 0 ) ...
                              + beta.*( D0.*CS_0 + DA.*CS_A + DB.*CS_B);
        elseif a>1 && b>1 && c>1
            CS(t,a,b,c,h,q) = ...
               log( (1/2)* ( 2 + exp( (v-pA(t,a,b,c,h,q))./mu ) ...
               + exp( (v-pB(t,a,b,c,h,q))./mu ) ) ) ...
```

```
            + beta.*( D0.*CS_0 + DA.*CS_A + DB.*CS_B);
    end

    % firm profits
    ProfA(t,a,b,c,h,q) = V1A(t,a,b,c,h,q)-kc*(a-1);
    ProfB(t,a,b,c,h,q) = V1B(t,a,b,c,h,q)-kc*(b-1);
    Profs(t,a,b,c,h,q) = ProfA(t,a,b,c,h,q) + ProfB(t,a,b,c,h,q);

end
```

check if global deviation

```
%candidates
V1A_cand = V1A(t,a,b,c,h,q);
V1B_cand = V1B(t,a,b,c,h,q);
pA_cand = pA(t,a,b,c,h,q);
pB_cand = pB(t,a,b,c,h,q);

% find global best
pA_glob_N = v*12*mm*(10/v)*2;
pB_glob_N = v*12*mm*(10/v)*2;
V1A_glob = zeros(pA_glob_N);
V1B_glob = zeros(pB_glob_N);

% e.g. if mm=5, run through all prices -12*v...+12*v, in steps of 2/v
for pA_glob_10 = 1:pA_glob_N
        pA_glob = pA_glob_10/(mm*(10/v)) - (pA_glob_N/2)/(mm*(10/v));
        [D0_glob,DA_glob,DB_glob] = ...
            demand_st_s(pA_glob,pB_cand,WW_C,a,b,h,q);
        V1A_glob(pA_glob_10) = DA_glob.*(pA_glob-mc) ...
            + beta.*(D0_glob.*W0_A+DA_glob.*WA_A+DB_glob.*WB_A);
end
for pB_glob_10 = 1:pB_glob_N
        pB_glob = pB_glob_10/(mm*(10/v)) - (pB_glob_N/2)/(mm*(10/v));
        [D0_glob,DA_glob,DB_glob] = ...
            demand_st_s(pA_cand,pB_glob,WW_C,a,b,h,q);
        V1B_glob(pB_glob_10) = DB_glob.*(pB_glob-mc) ...
            + beta.*(D0_glob.*W0_B+DA_glob.*WA_B+DB_glob.*WB_B);
end

pA_glob_best_10 = 1;
pB_glob_best_10 = 1;

for ppA_10 = 1:pA_glob_N
    if V1A_glob(ppA_10) >= V1A_glob(pA_glob_best_10)
        pA_glob_best_10 = ppA_10;
    end
end
for ppB_10 = 1:pB_glob_N
    if V1B_glob(ppB_10) >= V1B_glob(pB_glob_best_10)
        pB_glob_best_10 = ppB_10;
    end
end
```

```
pA_glob_best = pA_glob_best_10/(mm*(10/v)) - (pA_glob_N/2)/(mm*(10/v));
pB_glob_best = pB_glob_best_10/(mm*(10/v)) - (pB_glob_N/2)/(mm*(10/v));

%  Adjust pA and pB
if a == 1 || c==1
    pA_glob_best=realmax;
end
if b == 1 || c==1
    pB_glob_best=realmax;
end

% best VA
[D0_glob,DA_glob,DB_glob] = ...
    demand_st_s(pA_glob_best,pB_cand,WW_C,a,b,h,q);
V1A_glob_best = DA_glob.*(pA_glob_best-mc) ...
                + beta.*(D0_glob.*W0_A+DA_glob.*WA_A+DB_glob.*WB_A);
% best VB
[D0_glob,DA_glob,DB_glob] = ...
    demand_st_s(pA_cand,pB_glob_best,WW_C,a,b,h,q);
V1B_glob_best = DB_glob.*(pB_glob_best-mc) ...
                + beta.*(D0_glob.*W0_B+DA_glob.*WA_B+DB_glob.*WB_B);

% end condition for doneGlob
% first, assume all works
doneGlob = 1;
pA_0_G(t,a,b,c,h,q) = pA(t,a,b,c,h,q);
pB_0_G(t,a,b,c,h,q) = pB(t,a,b,c,h,q);

DifpA = abs(pA_glob_best-pA_cand);
DifpB = abs(pB_glob_best-pB_cand);

DifVA = (V1A_glob_best-V1A_cand);
DifVB = (V1B_glob_best-V1B_cand);

% done not 1 only if not global max
pAglobnot = 0;
pBglobnot = 0;
if ( (DifpA > ((v/20)/mm)) ...
        && (abs(V1A_glob_best - V1A_cand)>((v/20)/mm)) )
    disp(sprintf('pA NOT GLOBAL MAX'));
    doneGlob = 0;
    iterGlob = iterGlob + 1;
    pA_0_G(t,a,b,c,h,q) = pA_glob_best;
    pAglobnot = 1;
    if iterGlob==maxiter
        iterGlobexp(t,a,b,c,h,q) = 1;
    end
end
if ( (DifpB > ((v/20)/mm)) ...
        && (abs(V1B_glob_best - V1B_cand)>((v/20)/mm)) )
    disp(sprintf('pB NOT GLOBAL MAX'));
    doneGlob = 0;
    iterGlob = iterGlob + 1;
    pB_0_G(t,a,b,c,h,q) = pB_glob_best;
```

```
        pBglobnot = 1;
        if iterGlob==maxiter
            iterGlobexp(t,a,b,c,h,q) = 1;
        end
    end

    if (mod(iterGlob,10)==0) || done
        disp(sprintf('iterGlob=%d',iterGlob));
    end

    % if no EQ found, ok. Go to next starting value
    if iterGlob >= maxIterGlob
        doneGlob = 1 ;
        % take prices and value from previous starting value
        V1A(t,a,b,c,h,q) = V00A(t,a,b,c,h,q);
        V1B(t,a,b,c,h,q) = V00B(t,a,b,c,h,q);
        V1C(t,a,b,c,h,q) = V00C(t,a,b,c,h,q);
        pA(t,a,b,c,h,q) = pA00(t,a,b,c,h,q);
        pB(t,a,b,c,h,q) = pB00(t,a,b,c,h,q);
    end

end % end doneGlob

% end conditions for loop done 2
% Convergence in Values and prices and for done2 loop
tolVA = abs((V1A(t,a,b,c,h,q)-V00A(t,a,b,c,h,q))./ ...
    (1+abs(V1A(t,a,b,c,h,q))));
tolVB = abs((V1B(t,a,b,c,h,q)-V00B(t,a,b,c,h,q))./ ...
    (1+abs(V1B(t,a,b,c,h,q))));

tolPrA = abs((pA(t,a,b,c,h,q)-pA00(t,a,b,c,h,q))./ ...
    (1+abs(pA(t,a,b,c,h,q))));
tolPrB = abs((pB(t,a,b,c,h,q)-pB00(t,a,b,c,h,q))./ ...
    (1+abs(pB(t,a,b,c,h,q))));

% unique EQ = same EQ as with prior starting value
if ((tolPrA<tol) && (tolPrB<tol) && (tolVA<tol) && (tolVB<tol) ...
        || (iter2>=maxiter)) || ...
        (startval2==lowerstart && startval2==lowerstart)
    unique = 1;
else
    unique = 0;
    iter2 = iter2+1;
end

% if EQ not unique.. try again with updated starting values
if ((mod(iter2,10)==0) && iter2~=0 ) || done2
    disp(sprintf('No EQ'));
    pause
end

% Updating
V00A(t,a,b,c,h,q) = V1A(t,a,b,c,h,q);
V00B(t,a,b,c,h,q) = V1B(t,a,b,c,h,q);
```

```
V00C(t,a,b,c,h,q) = V1C(t,a,b,c,h,q);
pA00(t,a,b,c,h,q) = pA(t,a,b,c,h,q);
pB00(t,a,b,c,h,q) = pB(t,a,b,c,h,q);

if unique == 1
    startval2 = startval2+1;
    disp(sprintf('iter2=%d startval1=%g startval2=%g', ...
        iter2,startval1,startval2));
end
if unique == 1 && startval2 == upperstart && startval1~=upperstart
    startval1 = startval1+1;
    startval2 = lowerstart;
    disp(sprintf('iter2=%d startval1=%g startval2=%g', ...
        iter2,startval1,startval2));
end

% if for all starting values ok, finish done2 loop
if unique == 1 && startval2 == upperstart && startval1 == upperstart
    startval2 = lowerstart;
    startval1 = lowerstart;
    done2 = 1;
end

            end % end done2 loop

        end %q
        end %b
        end %a
        end %h
        end %c
    end % end for t<T

% end of computation
toc

end % end loop over all t

toc

% Print results.

diary off;
```

Published with MATLAB® R2014b

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%
% Dynamic Pricing under Capacity-constrained Competition with a Deadline
% Philipp Dimakopoulos, November 2017
%
% Set-up Function, Forward-Looking Consumers
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%


% Constants: Discounting.
r = 0.00;                           % discount rate.
beta = 1./(1+r);                    % discount factor.

% Constants: States. Starting Values.
T = 11;      % # time periods until good perishes
A = 16;      % # capacity A  % 1=>0
B = 16;      % # capacity B  % 1=>0
C = T;       % number of consumers waiting maximally in total %1=0 C>=T
H = 10;      % variations in mu
Q = 11;      % for variations of betac 0... 1

% Constants: Product market.
mc = 0;              % marginal costs
v = 100;             % quality of inside goods.
kc = 0;              % linear produciton costs of capacity
lowerstart = -20;    % starting value lower bound for uniqueness loop (*v/10)
upperstart = 20;     % starting value upper bound for uniqueness loop (*v/10)

% Constants: Program control.
tol = 1e-4;                 % tolerance.
maxiter = 5000;             % maximum number of iteration.
maxIterGlob = 25;           % max iters until: no EQ for these start values

% Multiplikator global deviation grid
mm = 5;

% Arrival Rate
arr = 1;

% Logit Demand. Consumer heterogeneity
mufix = v/10;
```

Published with MATLAB® R2014b

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%
% Dynamic Pricing under Capacity-constrained Competition with a Deadline
% Philipp Dimakopoulos, November 2017
%
% Beta c (patience) Function, Forward-Looking Consumers
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%


function [betacfu] = betacfun(q);

% Globals.
global Q;

% example
for i=1:Q
    if q == i
       betacfu = ((q-1)/(Q-1));
    end
end

% or fix betac!
%betacfu = 1;

end
```

Published with MATLAB® R2014b

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%
% Dynamic Pricing under Capacity-constrained Competition with a Deadline
% Philipp Dimakopoulos, November 2017
%
% Mu (Heterogeneity) Function, Forward-Looking Consumers
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

function [mufu] = mufun(h);

% Globals.
global beta v H mufix;

% example
for i=1:5
    if h == i
        mufu = (i)* ( (v/10) / 5 ); %2,4,6,8,10
    end
end
for i=6:7
    if h == i
        mufu = (i-3)* ( (v/10) / 2 ); %15,20
    end
end
for i=8:10
    if h == i
        mufu = (i-5)* ( (v/10) ); %30,40,50
    end
end

% or mufix
%mufu = mufix;

end
```

Published with MATLAB® R2014b

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%
% Dynamic Pricing under Capacity-constrained Competition with a Deadline
% Philipp Dimakopoulos, November 2017
%
% Demand Function, Forward-Looking Consumers
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

function [D0,D1,D2] = demand_st_s(p1,p2,WW_C,aa,bb,h,q);

% Globals.
global v;

% mu calculation
mu = mufun(h);
betac = betacfun(q);

% Demand.
uw = 2 * (exp( betac.*(WW_C) ) );

if aa == 1
    u1 = 0;
    uw = exp( betac.*(WW_C) );
else
    u1 = exp( ((v-p1)./mu) );
end
if bb == 1
    u2 = 0;
    uw = exp( betac.*(WW_C) );
else
    u2 = exp( ((v-p2)./mu) );
end

u = u1+u2+uw;

D0 = (uw./u);
D1 = (u1./u);
D2 = (u2./u);

end
```

Published with MATLAB® R2014b

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%
% Dynamic Pricing under Capacity-constrained Competition with a Deadline
% Philipp Dimakopoulos, November 2017
%
% First Order Condition A, Forward-Looking Consumers
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

function [Del,Jac] = FOC_st_s_A(p1,parms);

% Globals.
global beta H;

% Parameters: Marginal cost, rival's price, and expected values
mc1 = parms(1);
p2 = parms(2);
W0_A = parms(3);
WA_A = parms(4);
WB_A = parms(5);
WW_C = parms(6);
aa = parms(7);
bb = parms(8);
h = parms(9);
q = parms(10);

% mu calculation
mu = mufun(h);

% Demand.
[D0,D1,D2] = demand_st_s(p1,p2,WW_C,aa,bb,h,q);

% Objective.
Omega = D1.*(p1-mc1)+beta.*(D0.*W0_A+D1.*WA_A+D2.*WB_A);

% FOC.
Del = D1.*(1 - (1/mu).*(p1-mc1) - (1/mu).*beta.*WA_A + (1/mu).*Omega);

% Jacobian.
Jac = (1/mu).* ( Del.*(2.*D1 - 1)-D1 );
```

Published with MATLAB® R2014b

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%
% Dynamic Pricing under Capacity-constrained Competition with a Deadline
% Philipp Dimakopoulos, November 2017
%
% First Order Condition B, Forward-Looking Consumers
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

function [Del,Jac] = FOC_st_s_B(p2,parms);

% Globals.
global beta H;

% Parameters: Marginal cost, rival's price, and expected values.
mc2 = parms(1);
p1 = parms(2);
W0_B = parms(3);
WA_B = parms(4);
WB_B = parms(5);
WW_C = parms(6);
aa = parms(7);
bb = parms(8);
h = parms(9);
q = parms(10);

% mu calculation
mu = mufun(h);

% Demand.
[D0,D1,D2] = demand_st_s(p1,p2,WW_C,aa,bb,h,q);

% Objective.
Omega = D2.*(p2-mc2)+beta.*(D0.*W0_B+D1.*WA_B+D2.*WB_B);

% FOC.
Del = D2.*(1 - (1/mu).*(p2-mc2) - (1/mu).*beta.*WB_B + (1/mu).*Omega);

% Jacobian.
Jac = (1/mu) .* ( Del.*(2.*D2 - 1)-D2 );
```

Published with MATLAB® R2014b

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%
% Dynamic Pricing under Capacity-constrained Competition with a Deadline
% Philipp Dimakopoulos, November 2017
%
% First Order Condition AB, Forward-Looking Consumers
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%


function [Del,Jac] = FOC_st_s_AB(p,parms);

% Globals.
global beta H;

% Parameters: Marginal cost, rival's price, and expected values
mc_A = parms(1); mc_B = parms(2);
W0_A = parms(3); W0_B = parms(4);
WA_A = parms(5); WA_B = parms(6);
WB_A = parms(7); WB_B = parms(8);
WW_C = parms(9);
aa = parms(10); bb = parms(11);
h = parms(12);  q = parms(13);

% mu calculation
mu = mufun(h);

% Demand.
[D0,D1,D2] = demand_st_s(p(1),p(2),WW_C,aa,bb,h,q);

% Objective.
Omega = zeros(2,1);
Omega(1) = D1.*(p(1)-mc_A)+beta.*(D0.*W0_A+D1.*WA_A+D2.*WB_A);
Omega(2) = D2.*(p(2)-mc_B)+beta.*(D0.*W0_B+D1.*WA_B+D2.*WB_B);

% FOC.
Del = zeros(2,1);
Del(1) = D1.*(1-(1/mu).*(p(1)-mc_A)-(1/mu).*beta.*WA_A+(1/mu).*Omega(1));
Del(2) = D2.*(1-(1/mu).*(p(2)-mc_B)-(1/mu).*beta.*WB_B+(1/mu).*Omega(2));

% Jacobian.
Jac = zeros(2,2);
Jac(1,1) = (1/mu).* ( Del(1).*(2.*D1 - 1)-D1 );
Jac(1,2) = (1/mu).* D2.*Del(1) + (1/(mu^2)).*D1.*D2.*(Omega(1)-beta.*WB_A);
Jac(2,1) = (1/mu).* D1.*Del(2) + (1/(mu^2)).*D1.*D2.*(Omega(2)-beta.*WA_B);
Jac(2,2) = (1/mu).* ( Del(2).*(2.*D2 - 1)-D2 );
```

Published with MATLAB® R2014b

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%
% Dynamic Pricing under Capacity-constrained Competition with a Deadline
% Philipp Dimakopoulos, November 2017
%
% DogLeg (Root-Finding), Forward-Looking Consumers
% from Hans Bruun Nielsen,  IMM, DTU.  99.06.10
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

function  [X, info, perf] = DogLeg(fun,par, x0, opts)
%DogLeg  Dog Leg method for nonlinear system of equations
%    f_i(x) = 0 , i=1,...,n
%  where  x  is a vector,  x = [x_1, ..., x_n] .
%  In the discussion we also introduce the function
%    F(x) = .5 * sum(f_i(x)^2) .
%  The functions  f_i(x)  and the Jacobian matrix  J(x)  (with
%  elements  J(i,j) = Df_i/Dx_j ) must be given by a MATLAB
%  function with declaration
%            function  [f, J] = fun(x, par)
%  par  may be dummy.
%
%  Call:
%      [X, info {, perf}] = DogLeg(fun,par, x0, opts)
%
%  Input parameters
%  fun  :  String with the name of the function.
%  par  :  Parameters of the function.  May be empty.
%  x0   :  Starting guess for  x .
%  opts :  Vector with five elements:
%          opts(1) = Initial trust region radius.
%          opts(2:5) used in stopping criteria:
%              ||F'||inf <= opts(2)                     or
%              ||dx||2 <= opts(3)*(opts(3) + ||x||2)    or
%              ||f||inf <= opts(4)                      or
%              no. of iteration steps exceeds  opts(5) .
%
%  Output parameters
%  X    :  If  perf  is present, then array, holding the iterates
%          columnwise.  Otherwise, computed solution vector.
%  info :  Performance information, vector with 6 elements:
%          info(1:4) = final values of
%              [||f(x)||inf  ||F'||inf  ||dx||2  Delta]
%          info(5) = no. of iteration steps
%          info(6) = 1 :  Stopped by small  ||f(x)||inf
%                    2 :  Stopped by small  ||F'(x)||inf
%                    3 :  Stopped by small x-step
%                    4 :  Stopped by  kmax
%                    5 :  Problems, indicated by printout.
%  perf :  (optional). If present, then array, holding
%            perf(1,:) = values of  ||f(x)||inf
```

```
%            perf(2,:) = values of  ||F'(x)||inf
%            perf(3,:) = Radius of trust region,  Delta
%            perf(4,:) = values of  beta

%  Hans Bruun Nielsen,  IMM, DTU.  99.06.10

   %  Check and initialize
   [x n f J] = check(fun,par,x0,opts);
   thrc = max(20,n)*eps;     % For checking consistency
   g = J'*f;    ng = norm(g,inf);    ng2 = norm(g);    nf = norm(f,inf);
   delta = opts(1);    kmax = opts(5);
   F = (f'*f)/2;
   Trace = nargout > 2;
   if  Trace
         X = x*ones(1,kmax+1);
         perf = [nf; ng; delta; 0]*ones(1,kmax+1);
       end
   k = 1;    nu = 2;    stop = 0;    nx = opts(3) + norm(x);    beta = 0;
   nh = 0;    % added 04.05.12

   while  ~stop
     %  Check stopping criteria
     if       nf <= opts(4),   stop = 1;
     elseif  ng <= opts(2),   stop = 2;
     elseif  delta <= opts(3)*nx,   stop = 3;
     else     %  Find step
       alpha = (ng2/norm(J*g))^2;    a = -alpha*g;    na = alpha*ng2;
       [Q R] = qr(J);    y = Q'*(-f);
       D = abs(diag(R));
       si = find(D <= thrc*max(D));   nsi = length(si);
       if  nsi     % Singular.  Check consistency
         if  norm(y(si)) > thrc*F
           stop = 5;
           disp('Singular, non-consistent Newton equations.')
         else  % Find minimum norm solution
           p = ones(1,n);    p(si) = zeros(1,nsi);    p = find(p);
           RR = R(p,p);    b0 = [RR\y(p); zeros(nsi,1)];
           N = [RR\R(p,si); -eye(nsi)];
           b = b0 - N*(N\b0);
         end
       else,   b = R\y;   end
       if  ~stop     %   Proceed with Dog Leg
         nb = norm(b);
         if       nb <= delta     % Newton step
           h = b;    beta = 1;    nh = nb;    dL = F;
         elseif  na >= delta     %  Steepest descent
           h = -(delta/ng2)*g;    beta = 0;    nh = delta;
           dL = delta*(ng2 - .5*delta/alpha);
         else     % 'True' dog leg
           c = b - a;    cf = [c'*[c  2*a]  na^2-delta^2];
           beta = max(roots(cf));
           h = a + beta*c;    nh = delta;
           dL = .5*alpha*(1-beta)^2*ng2^2 + beta*(2-beta)*F;
         end
```

```
          if  nh <= opts(3)*nx,  stop = 3; end
        end
      end
      if  ~stop    % Perform step
        xnew = x + h;
        [fn,Jn] = feval(fun, xnew,par);   Fn = (fn'*fn)/2;
        dF = F - Fn;
        if  (dL > 0) & (dF > 0)
          x = xnew;   nx = opts(3) + norm(x);
          F = Fn;  J = Jn;  f = fn;  nf = norm(f,inf);
          g = J'*f;   ng = norm(g,inf);   ng2 = norm(g);
          delta = delta / max(1/3, (1 - (2*dF/dL - 1)^3));   nu = 2;
        else
          delta = delta / nu;  nu = 2*nu;
        end
        k = k + 1;
        if  Trace
              X(:,k) = x;
              perf(:,k) = [nf; ng; delta; 0];  perf(4,k-1) = beta;
            end
        if  k > kmax,  stop = 4; end
      end
    end
    %  Set return values
    if  Trace
      X = X(:,1:k);   perf = perf(:,1:k);
    else,  X = x;  end
    info = [nf  ng  nh  delta  k-1  stop];

% ==========  auxiliary function  =================================

function  [x,n, f,J] = check(fun,par,x0,opts)
%  Check function call
    sx = size(x0);   n = max(sx);
    if  (min(sx) > 1)
        error('x0  should be a vector'), end
    x = x0(:);   [f J] = feval(fun,x,par);
    sf = size(f);   sJ = size(J);
    if  sf(2) ~= 1 | sf(1) ~= n
        tx = 'f  must be a column vector of the same length as  x';
        error(tx), end
    if  sJ(1) ~= sf(1)
        error('row numbers in  f  and  J  do not match'), end
    if  any(sJ ~= n),  error('J  must be an n*n matrix'), end
%  Thresholds
    if  length(opts) < 5
        error('opts  must have 5 elements'), end
    if  length(find(opts(1:5) <= 0))
        error('The elements in  opts  must be strictly positive'), end
```

Published with MATLAB® R2014b

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%
% Dynamic Pricing under Capacity-constrained Competition with a Deadline
% Philipp Dimakopoulos, November 2017
%
% Transitiion Probability Function, Forward-Looking Consumers
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%



function [PrL_0,PrW_A,PrW_B] = transprobs_st_s(t,a,b,c,h,q);

% Globals.
global A B T C H Q arr;

% Transition probabilities.
PrW_A = zeros(T,A,B,C,H,Q);
PrW_B = zeros(T,A,B,C,H,Q);
PrL_0 = zeros(T,A,B,C,H,Q);

% ASS: choosing to buy at empty firm = waiting

% no consumer waiting in the market: if noone chose (0,A,B)
% relevant valuation matrix entry is prob of arrival
% note that capas unchanged as no consumer around to buy
if c==1
        % if arrival between t and t+1 -> c'=c+1
        PrL_0(min(T,t+1),a,b,min(C,c+1),h,q) = arr;
        PrW_A(min(T,t+1),a,b,min(C,c+1),h,q) = arr;
        PrW_B(min(T,t+1),a,b,min(C,c+1),h,q) = arr;

        % if no arrival between t and t+1 -> c'=c
        PrL_0(min(T,t+1),a,b,c,h,q) = 1-arr;
        PrW_A(min(T,t+1),a,b,c,h,q) = 1-arr;
        PrW_B(min(T,t+1),a,b,c,h,q) = 1-arr;

% at least one consumer, hence
% after arrival (and sale): c'=c-1+1=c
% after arrival (and no sale): c'=c+1=c+1
% after no arrival (and sale): c'=c-1
% after no arrival (and no sale): c'=c
else
    if a>1 && b>1
        PrL_0(min(T,t+1),a,b,min(C,c+1),h,q) = arr;
        PrW_A(min(T,t+1),a-1,b,c,h,q) = arr;
        PrW_B(min(T,t+1),a,b-1,c,h,q) = arr;

        PrL_0(min(T,t+1),a,b,c,h,q) = 1-arr;
        PrW_A(min(T,t+1),a-1,b,c-1,h,q) = 1-arr;
        PrW_B(min(T,t+1),a,b-1,c-1,h,q) = 1-arr;
```

```
    elseif a==1 && b>1
        PrL_0(min(T,t+1),a,b,min(C,c+1),h,q) = arr;
        PrW_A(min(T,t+1),a,b,min(C,c+1),h,q) = arr;
        PrW_B(min(T,t+1),a,b-1,c,h,q) = arr;

        PrL_0(min(T,t+1),a,b,c,h,q) = 1-arr;
        PrW_A(min(T,t+1),a,b,c,h,q) = 1-arr;
        PrW_B(min(T,t+1),a,b-1,c-1,h,q) = 1-arr;

    elseif a>1 && b==1
        PrL_0(min(T,t+1),a,b,min(C,c+1),h,q) = arr;
        PrW_A(min(T,t+1),a-1,b,c,h,q) = arr;
        PrW_B(min(T,t+1),a,b,min(C,c+1),h,q) = arr;

        PrL_0(min(T,t+1),a,b,c,h,q) = 1-arr;
        PrW_A(min(T,t+1),a-1,b,c-1,h,q) = 1-arr;
        PrW_B(min(T,t+1),a,b,c,h,q) = 1-arr;

    elseif a==1 && b==1
        PrL_0(min(T,t+1),a,b,min(C,c+1),h,q) = arr;
        PrW_A(min(T,t+1),a,b,min(C,c+1),h,q) = arr;
        PrW_B(min(T,t+1),a,b,min(C,c+1),h,q) = arr;

        PrL_0(min(T,t+1),a,b,c,h,q) = 1-arr;
        PrW_A(min(T,t+1),a,b,c,h,q) = 1-arr;
        PrW_B(min(T,t+1),a,b,c,h,q) = 1-arr;

    end

end


end
```

Published with MATLAB® R2014b

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%
% Dynamic Pricing under Capacity-constrained Competition with a Deadline
% Philipp Dimakopoulos, November 2017
%
% Lucky Consumer Transition Probability Func, Forward-Looking Consumers
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

function [PrC_0] = transprobs_st_s_cons(t,a,b,c,h,q,De0,DeA,DeB);

% Globals.
global A B T C H Q arr;

% Transition probabilities for waiting consumers.
PrC_0 = zeros(T,A,B,C,H,Q);
% prob to reach this state
prob = zeros(T,A,B,C,H,Q);
% end prob to be at this state (Zwischenstop)
eprob = zeros(T,A,B,C,H,Q);

% assumption: choosing to buy at empty firm = waiting
% end prob of current period = 1 (starting point)
eprob(t,a,b,c,h,1)=1;

for tt = t+1:+1:T-1 %nur bis T-1, denn in t=T V=0
    for aa = a:-1:1
    for bb = b:-1:1
    % prob and end prob of a state with no consumer (i.e. cc=1)
        % is not possible out of the perspective of a consumer
    for cc = 2:C
% I am the only one, i.e. in tt it will be my turn,
    % no matter what I choose to do
if cc==2
    % prob to reach this state is sum of
    % 1) after end prob tt-1 & no arrival,
    % not after arr since I am the only one and can not jsut have arrived
    prob(tt,aa,bb,cc,h,1) = 0
        + (1-arr)*( eprob(tt-1,aa,bb,cc,h,1) *( (cc-1)/(cc-1) ) );
    % end prob for this state is the sum of
    % 1) after the lucky consumer waited in tt [not possible cc=2]
    % 2) after endprob tt-1 & other cons did A or (3: B)
        % (i.e. I was unluncky cc-1/cc)
        % arr: in tt-1: aa/bb + 1, cc=cc+1-1,
        % in tt the other cons bought at aa/bb +1, cc=cc+1
        % no arr: in tt-1: aa/bb + 1, cc=cc+1,
        % in tt the other cons bought at aa/bb +1, cc=cc+1
    eprob(tt,aa,bb,cc,h,1) = 0 ...
       + arr *( eprob(tt-1,min(A,aa+1),bb,cc+1-1,h,1) ...
        *  DeA(tt,min(A,aa+1),bb,min(C,cc+1),h,1) * ( (cc-1)/cc ) )...
       + arr*( eprob(tt-1,aa,min(B,bb+1),cc+1-1,h,1) ...
```

```
            *  DeB(tt,aa,min(B,bb+1),min(C,cc+1),h,1) * ( (cc-1)/cc ) )...
           + (1-arr)*( eprob(tt-1,min(A,aa+1),bb,min(C,cc+1),h,1) ...
            *  DeA(tt,min(A,aa+1),bb,min(C,cc+1),h,1) * ( (cc-1)/cc ) )...
           + (1-arr)*( eprob(tt-1,aa,min(B,bb+1),min(C,cc+1),h,1) ...
            *  DeB(tt,aa,min(B,bb+1),min(C,cc+1),h,1) * ( (cc-1)/cc ) );
    else %cc>2 (at least two consumers)
        % Exclude possiility that after aa=1 someone could have chosen A
            % end prob for this state is the sum of
            % 1) after end prob in tt-1 & I am unlucky (cc-2/cc-1)
                % & in tt waiting of other cons
                    % arr: in tt-1: aa/bb, cc=cc-1,
                        % in tt the lucky cons waited at aa/bb, cc=cc
                    % no arr: in tt-1: aa/bb, cc=cc,
                        % in tt the lucky cons waited at aa/bb, cc=cc
            % 2) after endprob tt-1 & other cons did A or (3: B)
                 % I was unluncky before they bought,
                        % when we were 1 more cc: cc-1/cc
                    % arr: in tt-1: aa/bb + 1, cc=cc+1-1=cc,
                        % in tt the lucky cons bought at aa/bb +1, cc=cc+1
                    % no arr: in tt-1: aa/bb + 1, cc=cc+1,
                    % in tt the lucky cons bought at aa/bb +1, cc=cc+1
            eprob(tt,aa,bb,cc,h,1) = ...
             arr *( eprob(tt-1,aa,bb,max(1,cc-1),h,1) ...
                 * De0(tt,aa,bb,cc,h,1) * ( (cc-2)/(cc-1) ) )...
             + (1-arr)*( eprob(tt-1,aa,bb,cc,h,1) ...
                 * De0(tt,aa,bb,cc,h,1) * ( (cc-2)/(cc-1) ) )...
             + arr*( eprob(tt-1,min(A,aa+1),bb,cc,h,1) ...
                 * DeA(tt,min(A,aa+1),bb,min(C,cc+1),h,1) * ( (cc-1)/cc ) )...
             + (1-arr)*( eprob(tt-1,min(A,aa+1),bb,min(C,cc+1),h,1) ...
                 * DeA(tt,min(A,aa+1),bb,min(C,cc+1),h,1) * ( (cc-1)/cc ) )...
             + arr*( eprob(tt-1,aa,min(B,bb+1),cc,h,1) ...
                 *  DeB(tt,aa,min(B,bb+1),min(C,cc+1),h,1) * ( (cc-1)/cc ) )...
             + (1-arr)*( eprob(tt-1,aa,min(B,bb+1),min(C,cc+1),h,1) ...
                 * DeB(tt,aa,min(B,bb+1),min(C,cc+1),h,1) * ( (cc-1)/cc ) );
        end
        end
        end
        end
    end
    % in T use T-1 endprob as PrcW (V=0 anyways).
    % So ist Summe PrC_0 immer eins (zur Kontrolle)
    PrC_0 = prob;
    for aa = a:-1:1
        for bb = b:-1:1
        for cc = 1:C
            PrC_0(T,aa,bb,cc,h,1) = eprob(T-1,aa,bb,cc,h,1);
        end
        end
    end

    end
```

Published with MATLAB® R2014b

Bibliography

Abdulkadiroğlu, A. and Loertscher, S. (2007). Dynamic house allocation. *Working Paper*.

Abdulkadiroğlu, A., Pathak, P. A., and Roth, A. E. (2005a). The new york city high school match. *American Economic Review*, pages 364–367.

Abdulkadiroğlu, A., Pathak, P. A., Roth, A. E., and Sönmez, T. (2005b). The boston public school match. *American Economic Review*, pages 368–371.

Abdulkadiroğlu, A. and Sönmez, T. (2003). School choice: A mechanism design approach. *American Economic Review*, pages 729–747.

Allen, B. and Hellwig, M. (1986). Bertrand-edgeworth oligopoly in large markets. *The Review of Economic Studies*, 53(2):175–204.

Anderson, S. P. and Coate, S. (2005). Market provision of broadcasting: A welfare analysis. *The Review of Economic Studies*, 72(4):947–972.

Anderson, S. P. and Gabszewicz, J. J. (2006). The media and advertising: a tale of two-sided markets. *Handbook of the Economics of Art and Culture*, 1:567–614.

Armstrong, M. (2006). Competition in two-sided markets. *The RAND Journal of Economics*, 37(3):668–691.

Armstrong, M. and Wright, J. (2007). Two-sided markets, competitive bottlenecks and exclusive contracts. *Economic Theory*, 32(2):353–380.

Avery, C., Jolls, C., Posner, R., and Roth, A. E. (2007). The new market for federal judicial law clerks. pages 447–486.

Avery, C., Jolls, C., Posner, R. A., and Roth, A. E. (2001). The market for federal judicial law clerks. *The University of Chicago Law Review*, 68(3):793–902.

Aygün, O. and Bo, I. (2013). College admissions with multidimensional reserves: the brazillian affirmative action case. Technical report, Mimeo. 1.

Aygün, O. and Sönmez, T. (2012). The importance of irrelevance of rejected contracts in matchingn under weakened substitutes conditions. *Working Paper*.

Aygun, O. and Turhan, B. (2016). Dynamic reserves in matching markets: Theory and applications.

Balinski, M. and Sönmez, T. (1999). A tale of two mechanisms: Student placement. *Journal of Economic Theory*, 84(1):73–94.

Bloch, F. and Demange, G. (2017). Taxation and privacy protection on internet platforms. *Journal of Public Economic Theory*.

Bloch, F. and Houy, N. (2012). Optimal assignment of durable objects to successive agents. *Economic Theory*, 51(1):13–33.

Board, S. and Skrzypacz, A. (2016). Revenue management with forward looking buyers. *Journal of Political Economy*, 124(4):1046–1087.

Bonneau, J. and Preibusch, S. (2010). The privacy jungle: On the market for data protection in social networks. *Economics of information security and privacy*, pages 121–167.

Borenstein, S. and Rose, N. (1994). Competition and price dispersion in the us airline industry. *Journal of Political Economy*, 102(4).

Bourreau, M., Caillaud, B., and De Nijs, R. (2017). Taxation of a digital monopoly platform. *Journal of Public Economic Theory*. forthcoming.

Braun, S., Dwenger, N., and Kübler, D. (2010). Telling the truth may not pay off: An empirical study of centralized university admissions in germany. *The BE Journal of Economic Analysis & Policy*, 10(1).

Casadesus-Masanell, R. and Hervas-Drane, A. (2015). Competing with privacy. *Management Science*, 61(1):229–246.

Cooper, J. C. (2013). Privacy and antitrust: Underpants gnomes, the first amendment, and subjectivity. *George Mason Law & Economics Research Paper*, (13-39).

Crawford, V. P. and Knoer, E. M. (1981). Job matching with heterogeneous firms and workers. *Econometrica: Journal of the Econometric Society*, pages 437–450.

Crouzeix, J.-P. (1980). On second order conditions for quasiconvexity. *Mathematical Programming*, 18(1):349–352.

Damiano, E. and Lam, R. (2005). Stability in dynamic matching markets. *Games and Economic Behavior*, 52(1):34–53.

Dana Jr, J. D. (1999). Equilibrium price dispersion under demand uncertainty: the roles of costly capacity and market structure. *The RAND Journal of Economics*, pages 632–660.

De Corniere, A. and De Nijs, R. (2016). Online advertising and privacy. *The RAND Journal of Economics*, 47(1):48–72.

Debreu, G. (1952). A social equilibrium existence theorem. *Proceedings of the National Academy of Sciences*, 38(10):886–893.

Deneckere, R. and Peck, J. (2012). Dynamic competition with random demand and costless search: A theory of price posting. *Econometrica*, 80(3):1185–1247.

Dilme, F. and Li, F. (2016). Revenue management without commitment: dynamic pricing and periodic fire sales. *Available at SSRN 2435982*.

Dimakopoulos, P. D. and Heller, P. (2017). Matching with waiting times: The german entry-level labor market for lawyers. *mimeo*.

Dimakopoulos, P. D. and Sudaric, S. (2018). Privacy and platform competition. *International Journal of Industrial Organization*.

Doraszelski, U. and Satterthwaite, M. (2010). Computable markov-perfect industry dynamics. *The RAND Journal of Economics*, 41(2):215–243.

Dubins, L. E. and Freedman, D. A. (1981). Machiavelli and the gale-shapley algorithm. *American Mathematical Monthly*, pages 485–494.

Dudey, M. (1992). Dynamic edgeworth-bertrand competition. *The Quarterly Journal of Economics*, 107(4):1461–1477.

Dur, U. and Kesten, O. (2014). Sequential versus simultaneous assignment systems and two applications.

Edgeworth, F. Y. et al. (1925). Papers relating to political economy.

EDPS (2014). Privacy and competitiveness in the age of big data: The interplay between data protection, competition law and consumer protection in the digital economy. `https://edps.europa.eu/sites/edp/files/publication/14-03-26_competitition_law_big_data_en.pdf`. [Online; accessed 17-May-2017].

Escobari, D. (2012). Dynamic pricing, advance sales and aggregate demand learning in airlines. *The Journal of Industrial Economics*, 60(4):697–724.

European Commission (2015). Why we need a digital single market. `https://ec.europa.eu/commission/sites/beta-political/files/dsm-factsheet_en.pdf`. [Online; accessed 17-May-2017].

Fleiner, T. (2003). A fixed-point approach to stable matchings and some applications. *Mathematics of Operations Research*, 28(1):103–126.

Fudenberg, D. and Tirole, J. (1991). Game theory, 1991. *Cambridge, Massachusetts*, 393:12.

Gale, D. and Shapley, L. S. (1962). College admissions and the stability of marriage. *The American Mathematical Monthly*, 69(1):9–15.

Gallego, G. and Van Ryzin, G. (1994). Optimal dynamic pricing of inventories with stochastic demand over finite horizons. *Management Science*, 40(8):999–1020.

Gerardi, K. S. and Shapiro, A. H. (2009). Does competition reduce price dispersion? new evidence from the airline industry. *Journal of Political Economy*, 117(1):1–37.

Gönsch, J., Klein, R., Neugebauer, M., and Steinhardt, C. (2013). Dynamic pricing with strategic customers. *Journal of Business Economics*, 83(5):505–549.

Hanseatisches Oberlandesgericht (2015). Gewichtete bewerbungsliste.

Hansestadt Hamburg (2012). Verordnung über die aufnahme in den juristischen vorbereitungsdienst (aufnahmeverordnung) vom 27. januar 2004.

Haruvy, E., Roth, A. E., and Ünver, M. U. (2006). The dynamics of law clerk matching: An experimental and computational investigation of proposals for reform of the market. *Journal of Economic Dynamics and Control*, 30(3):457–486.

Hatfield, J. W. and Kojima, F. (2008). Matching with contracts: Comment. *The American Economic Review*, 98(3):1189–1194.

Hatfield, J. W. and Kojima, F. (2009). Group incentive compatibility for matching with contracts. *Games and Economic Behavior*, 67(5):745–749.

Hatfield, J. W. and Kojima, F. (2010). Substitutes and stability for matching with contracts. *Journal of Economic Theory*, 145(5):1704–1723.

Hatfield, J. W., Kominers, S. D., Nichifor, A., Ostrovsky, M., and Westkamp, A. (2015a). Full substitutability in trading networks. In *Proceedings of the Sixteenth ACM Conference on Economics and Computation*, pages 39–40. ACM.

Hatfield, J. W., Kominers, S. D., and Westkamp, A. (2015b). Stability, strategy-proofness, and cumulative offer mechanisms. *Working Paper.*

Hatfield, J. W. and Milgrom, P. R. (2005). Matching with contracts. *American Economic Review*, pages 913–935.

Hörner, J. and Samuelson, L. (2011). Managing strategic buyers. *Journal of Political Economy*, 119(3):379–425.

Justizprüfungsamt Hessen (2011). Merkblatt für bewerberinnen und bewerber um einstellung in den juristischen vorbereitungsdienst in hessen.

Kadam, S. V. and Kotowski, M. H. (2015). Multi-period matching. *Working Paper.*

Katzner, D. W. et al. (1970). Static demand theory.

Kelso, A. S. and Crawford, V. P. (1982). Job matching, coalition formation, and gross substitutes. *Econometrica*, pages 1483–1504.

Kennes, J., Monte, D., and Tumennasan, N. (2014). The day care assignment: A dynamic matching problem. *American Economic Journal: Microeconomics*, 6(4):362–406.

Kominers, S. D. and Sönmez, T. (2016). Matching with slot-specific priorities: Theory. *Theoretical Economics*, 11(2):683–710.

Kreps, D. M. and Scheinkman, J. A. (1983). Quantity precommitment and bertrand competition yield cournot outcomes. *The Bell Journal of Economics*, pages 326–337.

Kummer, M. and Schulte, P. (2016). When private information settles the bill.

Kurino, M. (2009). Credibility, efficiency and stability: A theory of dynamic matching markets. *Jena Economic Research Papers, JENA.*

Kurino, M. (2014). House allocation with overlapping generations. *American Economic Journal: Microeconomics*, 6(1):258–89.

Land Brandenburg (2014). Juristenausbildungsgesetz vom 4. juni 2004.

Laudon, K. C. (1996). Markets and privacy. *Communications of the ACM*, 39(9):92–104.

Lefouili, Y. and Toh, Y. L. (2017). Privacy and quality. *TSE Working Paper*, 17(795).

Leshno, J. D. (2015). Dynamic matching in overloaded waiting lists. *Working Paper*.

Levitan, R. and Shubik, M. (1972). Price duopoly and capacity constraints. *International Economic Review*, pages 111–122.

Lin, K. Y. and Sibdari, S. Y. (2009). Dynamic price competition with discrete customer choices. *European Journal of Operational Research*, 197(3):969–980.

Martínez-de Albéniz, V. and Talluri, K. (2011). Dynamic price competition with fixed capacities. *Management Science*, 57(6):1078–1093.

McFadden, D. et al. (1973). Conditional logit analysis of qualitative choice behavior.

Meisner, V. (2017). Competing for strategic buyers. *mimeo*.

Möller, M. and Watanabe, M. (2010). Advance purchase discounts versus clearance sales*. *The Economic Journal*, 120(547):1125–1148.

Niederle, M. and Roth, A. E. (2003). Unraveling reduces mobility in a labor market: Gastroenterology with and without a centralized match. *Journal of Political Economy*, 111(6):1342–1352.

Niederle, M. and Roth, A. E. (2009). Market culture: How norms governing exploding offers affect market performance. *American Economic Journal: Microeconomics*, 1:199–219.

Nocke, V. and Peitz, M. (2007). A theory of clearance sales*. *The Economic Journal*, 117(522):964–990.

Oberlandesgericht Düsseldorf (2015a). Gesuch um aufnahme in den juristischen vorbereitungsdienst.

Oberlandesgericht Düsseldorf (2015b). Merkblatt zum gesuch um aufnahme in den juristischen vorbereitungsdienst.

Oberlandesgericht München, (2015). Kriterien für die zuweisung der ausbildungsplätze.

Pakes, A. and McGuire, P. (1992). Computing markov perfect nash equilibria: Numerical implications of a dynamic differentiated product model.

Pathak, P. A. and Sönmez, T. (2013). School admissions reform in chicago and england: Comparing mechanisms by their vulnerability to manipulation. *American Economic Review*, 103(1):80–106.

Peitz, M. and Reisinger, M. (2016). The economics of internet media. In *Handbook of Media Economics, vol. 1A*. Elsevier.

Peitz, M. and Valletti, T. M. (2008). Content and advertising in the media: Pay-tv versus free-to-air. *International Journal of Industrial Organization*, 26(4):949–965.

Perakis, G. and Sood, A. (2006). Competitive multi-period pricing for perishable products: A robust optimization approach. *Mathematical Programming*, 107(1-2):295–335.

Pew Research Center (2014). Public perceptions of privacy and security in the post-snowden era. `http://www.pewinternet.org/files/2014/11/PI_PublicPerceptionsofPrivacy_111214.pdf`. [Online; accessed 17-May-2017].

Posner, R. A. (1981). The economics of privacy. *The American Economic Review*, 71(2):405–409.

Reisinger, M. (2012). Platform competition for advertisers and users in media markets. *International Journal of Industrial Organization*, 30(2):243–252.

Roth, A. E. (1982). The economics of matching: Stability and incentives. *Mathematics of Operations Research*, 7(4):617–628.

Roth, A. E. (1984). The evolution of the labor market for medical interns and residents: a case study in game theory. *The Journal of Political Economy*, 92(6):991.

Roth, A. E. (1991). A natural experiment in the organization of entry-level labor markets: regional markets for new physicians and surgeons in the united kingdom. *The American Economic Review*, pages 415–440.

Roth, A. E. and Peranson, E. (1999). The redesign of the matching market for american physicians: Some engineering aspects of economic design. *American Economic Review*, 89(4):748–779.

Roth, A. E., Sönmez, T., and Ünver, M. U. (2004). Kidney exchange. *The Quarterly Journal of Economics*, pages 457–488.

Roth, A. E. and Sotomayor, M. A. O. (1990). *Two-sided matching: A study in Game-theoretic Modeling and Analysis.* Number 18. Cambridge University Press.

Roth, A. E. and Xing, X. (1994). Jumping the gun: Imperfections and institutions related to the timing of market transactions. *The American Economic Review*, pages 992–1044.

Roth, A. E. and Xing, X. (1997). Turnaround time and bottlenecks in market clearing: Decentralized matching in the market for clinical psychologists. *Journal of Political Economy*, 105(2):284–329.

Schummer, J. and Abizada, A. (2015). Incentives in landing slot problems. *Working Paper*.

Schummer, J. and Vohra, R. V. (2013). Assignment of arrival slots. *American Economic Journal: Microeconomics*, 5(2):164–185.

Sinitsyn, M. (2008). Characterization of the support of the mixed strategy price equilibria in oligopolies with heterogeneous consumers. *Economics Letters*, 99(2):242–245.

Sönmez, T. (2013). Bidding for army career specialties: Improving the rotc branching mechanism. *Journal of Political Economy*, 121(1):186–219.

Sönmez, T. and Switzer, T. B. (2013). Matching with (branch-of-choice) contracts at the united states military academy. *Econometrica*, 81(2):451–488.

Spiegel, Y. (2013). Commercial software, adware, and consumer privacy. *International Journal of Industrial Organization*, 31(6):702–713.

Stadt Berlin (2004). Verordnung über die ausbildungskapazität und das vergabeverfahren für den juristischen vorbereitungsdienst (berlin kapvvo).

Stucke, M. E. and Grunes, A. P. (2016). *Big Data and Competition Policy*. Oxford University Press.

Su, X. (2007). Intertemporal pricing with strategic customer behavior. *Management Science*, 53(5):726–741.

Sweeting, A. (2012). Dynamic pricing behavior in perishable goods markets: Evidence from secondary markets for major league baseball tickets. *Journal of Political Economy*, 120(6):1133–1172.

Talluri, K. T. and Van Ryzin, G. J. (2006). *The theory and practice of revenue management*, volume 68. Springer Science & Business Media.

Thakral, N. (2015). Matching with stochastic arrival. *Working Paper*.

Train, K. E. (2009). *Discrete choice methods with simulation*. Cambridge University Press.

Vives, X. (2001). *Oligopoly pricing: old ideas and new tools*. MIT press.

Waehrer, K. (2015). Online services and the analysis of competitive merger effects in privacy protections and other quality dimensions. *Available at SSRN*.

Westkamp, A. (2013). An analysis of the german university admissions system. *Economic Theory*, 53(3):561–589.

Williams, K. R. (2017). Dynamic airline pricing and seat availability.

Selbständigkeitserklärung

Ich bezeuge durch meine Unterschrift, dass meine Angaben über die bei der Abfassung meiner Dissertation benutzten Hilfsmittel, über die mir zuteil gewordene Hilfe sowie über frühere Begutachtungen meiner Dissertation in jeder Hinsicht der Wahrheit entsprechen.

Berlin, 21. Dezember 2017

Philipp D. Dimakopoulos